

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 09-05-03
Searcher: Baldwin 2444
Terminal time: 20
Elapsed time: _____
CPU time: _____
Total time: 23
Number of Searches: _____
Number of Databases: 1

Search Site
____ STIC
____ CM-1
____ Pre-S
Type of Search
____ N.A. Sequence
____ A.A. Sequence
____ Structure
____ Bibliographic

Vendors
____ IG
____ STN
____ Dialog
____ APS
____ Geninfo
____ SDC
____ DARC/Questel
____ Other CGN

BEST AVAILABLE COPY

Fri Sep 5 09:48:11 2003

us-09-937-982-3.inpm

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2003, 16:24:21; Search time 3324 Seconds

(without alignments)
11169.064 Million cell updates/sec

Title: US-09-937-982-3

Perfect score: 1133

Sequence: 1 atggcagcagcatcatca.....aagaaggaatctcttcac 1133

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 32997241 seqs, 16383922548 residues 65994482

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Pending_Patents_MA.Main.*
1: /cgn2_6/ptodata/1/pna/US0937A.COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US0938A.COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US0939A.COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US0940A.COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US0941A.COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US0942A.COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US0943A.COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US0944A.COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US0945A.COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US0946A.COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US0947A.COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US0948A.COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US0949A.COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US0950A.COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US0951A.COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US0952A.COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US0953A.COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US0954A.COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US0955A.COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US0956A.COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US0957A.COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US0958A.COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US0959A.COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US0960A.COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US0961A.COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US0962A.COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US0963A.COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US0964A.COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US0965A.COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US0966A.COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US0967A.COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US0968A.COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US0969A.COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US0970A.COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US0971A.COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US0972A.COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US0973A.COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US0974A.COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US0975A.COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US0976A.COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US0977A.COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US0978A.COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US0979A.COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1133	100.0	1133	39	US-09-937-982-3
2	1129.8	99.7	1133	33	US-09-773-260-2
3	302.8	26.7	9082	37	US-09-897-516-1293
4	302.8	26.7	9082	37	US-09-897-516-1294

WALICKA
09/19/98
Seq. ID 3 w/ Interf
Page 1

GENERAL INFORMATION:

APPLICANT: Corbin, David R.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Huesing, Joseph E.
 APPLICANT: Krasomil-Osterfeld, Karina C.
 APPLICANT: Malvar, Thomas M.
 APPLICANT: Slater, Steven C.
 APPLICANT: Spilidonov, Sergei
 TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 FILE REFERENCE: 38-21(51847)B
 CURRENT APPLICATION NUMBER: US/09/897,516
 PRIOR FILING DATE: 2001-06-29
 PRIOR APPLICATION NUMBER: US 60/215, 161
 NUMBER OF SEQ ID NOS: 8409
 SEQ ID NO 1296
 LENGTH: 9082
 TYPE: DNA
 ORGANISM: Xenorhabdus sp.
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (8058)..(9080)
 US-09-897-516-1296

Query Match 26.7%; Score 302.8; DB 37; Length 9082;
 Best Local Similarity 58.1%; Pred. No. 3.4e-81;

Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

127 CCCCAGTGTACTTCTAGCCACAGAGCACCATCCCTGGTGGGGAATCTAGAGTGC 186
 5328 CCAACATTAATGCTTGGCAGAGGCGGCACATCCCTGGTGGTGAATCTGCAACT 5387
 187 AAGAGTGTACTTCTGTGGAGCAGTACCGGTGATAGCTTCTTGGAGCCCTGCC 246
 5388 AATCCAGCTACCTGCGGCTCAAGTGGTGGTATCATGATGATGATGATGATGATG 5447
 247 ATCAAGACCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 306
 5448 ATCAAAACATTTCTATCTGAAAGGAGACAGGTTTATGCAATCGGTTCTCAGACATG 5507
 307 ACCGGTAAGTGTGGCTTAAATAGCCAGCAGTGTCAATAGCTCTCCGCCAAAAGAG 366
 5508 AACGATCAGGTCTGGCTGAGTCTGGCGAAAAAATCAATG-----CAGATTGTGATATA 5561
 367 ACCGAGCCGTGATCATACCCAGTACAGTACATCATGAGAGAGAGCCCTTCTCTC 426
 5562 ACTGACGGTTTGTCTATCACCACGACGATACCATGGAAGAAACTGCTTATTTCTC 5621
 427 AACCTCAGGTGAAAGCCAAACCTGCTGCTTGTAGGCGCCATGCGTCCAGGCTC 486
 5622 GATCTGACCATGATGTATTAACCTATGCTATGTTGGCCCATGCTCCATCAAG 5681
 487 TCCATAGTGTGATGAGCCCATGAAATCTCTATACGCGGTGATAGGATCAACAAA 546
 5682 GCTGTGGGTGATGAGCCCATGAACTCTATTAATGCTGTTGTAGCGTCAAGTAAA 5741
 547 GCTCTACTAACAAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATG 606
 5742 GACTGTGAAACCGTGTGTTTACTTGTATGAAATGATGATGATGATGATGATGATG 5801
 607 GCGACCAAGCTCAACACACCGAGTCAATGATTTGCTTGGCCCAACAGATTAATC 666
 5802 ATCGGTAACAGTACAGCACTGAAAGTTTCAAGGCTTTCAGGACAGTGAAGAGCCACAA 5861
 667 GGCACAGTCTATTATGGAAGTCAAGTCTTCTACATCAATCGTTGACCTCAGACCTT 726
 5862 GCTTCATTCATTAACGTAAGTAAGTAAGTACTCTCT--GCAGCACAAGTGAAGAGAT 5918
 727 GCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 786
 5919 AAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5978

RESULT 7

US-09-897-516-1297

Sequence 1297, Application US/09897516

GENERAL INFORMATION:

APPLICANT: Corbin, David R.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Huesing, Joseph E.
 APPLICANT: Krasomil-Osterfeld, Karina C.
 APPLICANT: Malvar, Thomas M.
 APPLICANT: Slater, Steven C.
 APPLICANT: Spilidonov, Sergei
 TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 FILE REFERENCE: 38-21(51847)B
 CURRENT APPLICATION NUMBER: US/09/897,516
 PRIOR FILING DATE: 2001-06-29
 PRIOR APPLICATION NUMBER: US 60/215, 161
 NUMBER OF SEQ ID NOS: 8409
 SEQ ID NO 1297
 LENGTH: 9082
 TYPE: DNA
 ORGANISM: Xenorhabdus sp.
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (6861)..(7958)
 US-09-897-516-1297

Query Match 26.7%; Score 302.8; DB 37; Length 9082;
 Best Local Similarity 58.1%; Pred. No. 3.4e-81;
 Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

127 CCCCAGTGTACTTCTAGCCACAGAGCACCATCCCTGGTGGGGAATCTAGAGTGC 186
 5328 CCAACATTAATGCTTGGCAGAGGCGGCACATCCCTGGTGGTGAATCTGCAACT 5387
 187 AAGAGTGTACTTCTGTGGAGCAGTACCGGTGATAGCTTCTTCCAGCCCTGCC 246
 5388 AATCCAGCTACCTGCGGCTCAAGTGGTGGTATTCATTTGCTGATGATGATGATG 5447
 247 ATCAAGACCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 306
 5448 ATCAAAACATTTCTATCTGAAAGGAGACAGGTTTATGCAATCGGTTCTCAGACATG 5507
 307 ACCGGTAAGTGTGGCTTAAATAGCCAGCAGTGTCAATAGCTCTCCGCCAAAAGAG 366
 5508 AACGATCAGGTCTGGCTGAGTCTGGCGAAAAAATCAATG-----CAGATTGTGATATA 5561

QY	367	ACGGAAGCCGTAATCAATACACCCATGGAAATCGAACAATGGAGAAGACCGCTTCTCTC	426
Db	5562	ACGACGCGTTTGTCTATCACCACCGGTACAGATACATGGAGAAGAACTGCTTATTTCTC	5621
QY	427	AACCTCAAGGTGAAAGAACCCAAAAACCTGCTGCTGCTGTAGCGCCATGCGCTCCAGGCTCT	486
Db	5622	GATCTGACCACTAGTGTCTATAAACCTATGATATGATGTTGGCCGAATGCTTCATCAACG	5681
QY	487	TCCATGAGTGTATGAGCCCATGAAATCTCTATTAACCCCTGAAATGTAAGGATCAACAA	546
Db	5682	GCTCTGGGTGTGATGGCCCATTTGAACCTCTATATATCTGTGTGTAGAGTCAGATAAA	5741
QY	547	GCTCTACTATAACAAGAGATGATGATGTATGATGAACATGAGATTCAAGCCGCGAGAA	606
Db	5742	GACTCTGCMAAACCTGGTGTGTTTACTGTGATGATGATATCTGTATTCATTAAGACGTAC	5801
QY	607	GGCAACAACTCAACACCAACCGCAGTAAATGATTTGCTTGCCTCAACACAGTAATATC	666
Db	5802	ATCGGTAACTGACACCAACATGAAAGTTTCAGGCTTCCAGGATGAGAGGAGGCCCAAA	5861
QY	667	GGCAAGCTCTTATATGCAAAAGTGAATTTACTCAATCCGTTGCACTCAACACCTT	726
Db	5862	GCTTCTATCTATTAACGGTAAGTAATACACTACTCT---GAGACACAGTAAGAACGAT	5918
QY	727	GCAAGTAGTTGATATTAGCAAAATCGAAGACTCCCGAGAGTGCATTTCTTACGCT	786
Db	5919	AAAGGTGTTTTTGTATGTCAGCAAACTGCAACGTGCCAAAAGTAGGATTTGTATTAAC	5978
QY	787	CACCCCGATGATACGATGATTTTATGTCATGACCCCTTCAGGAGAGGCCAAAGATC	846
Db	5979	TACTCTAATGCTTCGCACTCTCCCGGTAAACATTTGTAGAGATGATATTAAGATAT	6038
QY	847	ATCCATGACGAGCATGGGCATGGGAACCTTTCCCTTTGACTCAAAATGCTCTGAAAAA	906
Db	6039	ATCAATGCTGGTGTGTGAAGGTAACATCTATAAGCTATTCTGCACTCTCTCAAA	6098
QY	907	GCAGCAATACAGGCGTATGCTGCTGCTGCGTGAAGCTCTGAGTGGGCGAGTGGTCCACACC	966
Db	6099	GCGGTAAAGATGGTGTGTGTGTGTGTGTTCCAACTATTTCTTTCGTTTCACTACT	6158
QY	967	CAGAAGCGTGAAGTGATGATTAAGAAACTTGTGTTGTGGCTACAGAGATCTCAACCT	1026
Db	6159	CAAAATGCTGAAGTTAATGACAGCAAAATAGCGTTTGTTCATCAGAAAGCCTGAACCG	6218
QY	1027	CAAAAAGCCAGAGTCTCTTATGTATGTTAGCCCTCACCACAAACTAGTATGATAGAGAGCGATC	1086
Db	6219	CAGAAAGCTCGTCTTCTTCTGCAATTTGCTCTGACTGAGCTGCTGACACAGCAAAATTT	6278
QY	1087	CAAAAGATCTCTCCACCTATTATCAAGAA	1118
Db	6279	CAGGAATGTTTTCTTAATATTAAATTCCTGAA	6310
RESULT 8			
US-09-897-516-1298			
Sequence 1298, Application US/09897516			
GENERAL INFORMATION:			
APPLICANT: Corbin, David R.			
APPLICANT: Goldman, Barry S.			
APPLICANT: Hinkle, Gregory J.			
APPLICANT: Huesing, Joseph E.			
APPLICANT: Krasomil-Osterfeld, Karina C.			
APPLICANT: Malvar, Thomas M.			
APPLICANT: Slater, Steven C.			
APPLICANT: Spiridonov, Sergei			
TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof			
FILE REFERENCE: 38-21(5184)B			
CURRENT APPLICATION NUMBER: US/09/897,516			
CURRENT FILING DATE: 2001-06-29			
PRIOR APPLICATION NUMBER: US 60/715, 161			
PRIOR FILING DATE: 2000-06-30			
NUMBER OF SEQ ID NOS: 8409			
SEQ ID NO 1298			

[illegible]

Db 6219 CAGAAAGCTGCTGTTCTTCTGCAATGCTGCACTGAGTGGTGAACAGCAAAAT 6278
 Oy 1087 CAAAAGATCTTCTCCACCTATTAATCCAGAA 1118
 Db 6279 CAGAAATGTTTCTTAATATTAATTCCTGAA 6310

RESULT 9 US-60-215-161-1293/c

Sequence 1293, Application US/60215161
 GENERAL INFORMATION:
 APPLICANT: Corbin, David R.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Huesling, Joseph E.
 APPLICANT: Krasomil-Osterfeld, Karina C.
 APPLICANT: Malvar, Thomas M.
 APPLICANT: Slater, Steven C.
 APPLICANT: Spiridonov, Sergei
 APPLICANT: Hinkle, Gregory J.
 TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 FILE REFERENCE: 38-21(51847)A
 CURRENT APPLICATION NUMBER: US/60/215,161
 CURRENT FILING DATE: 2000-06-30
 NUMBER OF SEQ ID NOS: 8409
 SEQ ID NO 1293
 LENGTH: 9082
 TYPE: DNA
 ORGANISM: Xenorhabdus sp.
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (3868)..(5734)
 US-60-215-161-1293

Query Match 26.7%; Score 302.8; DB 75; Length 9082;
 Best Local Similarity 58.1%; Pred. No. 3.4e-81;
 Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

Oy 127 CCCCAAGTACTTCTTCTAGCCACAGAGGACCATCGTGGGGAATCTAGGCTC 186
 Db 3755 CCAAAATTAATCTTCTGCGACGGGTGGCACAATCGTGGTGAATCGCAACT 3696
 Oy 187 AAGAGTACTTCTGCTGAGCAGTACCGTGAATGCTTCTGAGCGGCCCTGCC 246
 Db 3695 AATTCAGCTACACTGCGGTGCAAGTGGTGTGATTCATTGCTGAATGCTGTCGCT 3636
 Oy 247 ATACAGACTTACCCACCATCAAGGGTGAACAGATCTCAAGTCTGCCCAAGAGAT 306
 Db 3635 ATCAAAAACATCTTAATCTGAAGAGTGACAGATTGTAGCATCGGTTCTCAGAGATG 3576
 Oy 307 ACGGTAAGGTGGGTTAACTAGCCACAGCGTGTCAATGAGTCCCTGCCCAAAAAGAG 366
 Db 3575 AACGATCAGGTCTGGCTGACTGCGCAAAAATCAATG-----CAGATTGTGATAA 3522
 Oy 367 ACCGAAGCCGTGATCAATCAACCATGAGACCATGAGAGAGACCGCTTCTTCTC 426
 Db 3521 ACTGACGGTTTGTCTATCAACCGGTAGACGATACCATGAGAAACTGTTATTCTC 3462
 Oy 427 AACCTACGCTGAAAAACCAAACTGTCTGCTTGTAGCGCCATGCGTCCAGGCTC 486
 Db 3461 GATCTGACACCTGAGTGTATTAACCTATGTAATGTTGGCCCAATGCGTCCATCAAG 3402
 Oy 487 TCCATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
 Db 3401 GCTCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3342
 Oy 547 GCTCTACTACAAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 606
 Db 3341 GACTCTGCAACCGTGTGTTTACTTGTATGAATGATTTCTATTTATGATGAGCTGAC 3282
 Oy 607 GGGACCAACTCAACCAACCGAGTCAATGCTTGTGCTTCCGCCAAGCAGGTAATC 666
 Db 3281 ATCGGTAAGTACGACCAACTGAATGCTTGTGCTTCCAGGAGTGAACGAGGCCACAA 3222

Oy 667 GGCACGTCTATTAATGCAAAAGTGCATATTTCACTCAATCCGTTGACCTCACACCTT 726
 Db 3221 GCTTCATCTATTAATGCAAAAGTGAATTAATTAATCTACTCT---GGAGCACAGTGAACAGAT 3165
 Oy 727 GCAAGTGTGTTGATTAATGCAAAAGTGAATTAATTAATTAATTAATTAATTAATTAAT 786
 Db 3164 AAGCTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3105
 Oy 787 CACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
 Db 3104 TACTCTAATGCTTCCGATCTGCGGTGAACGATTTGATGATGATGATGATGATGATGAT 3045
 Oy 847 ATCCATGAGCAGTGGGCAATGGAACCTTTCCCTTGTGACTCAAAATGCTCTGAAAA 906
 Db 3044 ATCAGTGTGTTGTTGTAACGGAATCAATTAATTAATTAATTAATTAATTAATTAAT 2985
 Oy 907 GCAGCAATCAGCGTGTGCTGCTGAGCTCTAGTGGGCAAGTGGTTCACACCC 966
 Db 2984 GCGGCTAAGATGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2925
 Oy 967 CAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026
 Db 2924 CAAATGCTGAATTAATGACGCAAAATACGTTTGTGATGATGATGATGATGATGATGAT 2865
 Oy 1027 CAAAAGCAGAGCTTCTATGTTAGCCCTCACCAAAATAGTATGATGATGATGATGATGAT 1086
 Db 2864 CAGAAAGCTGCTGTTCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2805
 Oy 1087 CAAAAGATCTTCTCACCCTATTAATCCAGAA 1118
 Db 2804 CAGGAATGTTTCTTAATATTAATTCCTGAA 2773

RESULT 10

US-60-215-161-1294/c
 Sequence 1294, Application US/60215161

GENERAL INFORMATION:
 APPLICANT: Corbin, David R.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Huesling, Joseph E.
 APPLICANT: Krasomil-Osterfeld, Karina C.
 APPLICANT: Malvar, Thomas M.
 APPLICANT: Slater, Steven C.
 APPLICANT: Spiridonov, Sergei
 APPLICANT: Hinkle, Gregory J.
 TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 FILE REFERENCE: 38-21(51847)A
 CURRENT APPLICATION NUMBER: US/60/215,161
 CURRENT FILING DATE: 2000-06-30
 NUMBER OF SEQ ID NOS: 8409
 SEQ ID NO 1294
 LENGTH: 9082
 TYPE: DNA
 ORGANISM: Xenorhabdus sp.
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (6073)..(6924)
 US-60-215-161-1294

Query Match 26.7%; Score 302.8; DB 75; Length 9082;
 Best Local Similarity 58.1%; Pred. No. 3.4e-81;
 Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

Oy 127 CCCCAAGTACTTCTTCTAGCCACAGAGGACCATCGTGGGGAATCTAGGCTC 186
 Db 3755 CCAAAATTAATCTTCTGCGACGGGTGGCACAATCGTGGTGAATCGCAACT 3696
 Oy 187 AAGAGTACTTCTGCTGAGCAGTACCGTGAATGCTTCTGAGCGGCCCTGCC 246
 Db 3695 AATTCAGCTACACTGCGGTGCAAGTGGTGTGATTCATTGCTGAATGCTGTTCTGCT 3636
 Oy 247 ATCAAGCACTTACCCACCATCAAGGTTGAACAGATCTCAAGCATTTGGTCCCAAGAGATG 306

```

Db      3635 ATCAAAACATTGTAATCTGAAAGTGAAGCGAGTTGTTAGCATGGTTCTCAGACATG 3576
      307  ACGGGTAAGTGTGGCTTAATTAAGTACCAAGCGTGTCAATGAGCTCCGCCCAAAAGAG 366
      3575 AACGATCAGTGTGGCTGACCTGTGGCAAAAAATCAATG-----CAGATTGATATA 3522
      367  ACCGAAGCCGTGATCATCAACCATGAACTGACACCATGGAAGAGAACCGCTTCTCTC 426
      3521 ACTGACGGTTTTGTGATCACCCACGCTACAGATACAGGAAGAAAGCTATTCTCTC 3462
      427  AACCTCAGGGTAAAAAGCCAAACCTGCTCTGCTAGGCGCATCGCTCAGGCTCT 486
      3461 GATCTGACACCTCAGTGTATTAACCTATGTAATGTTGGGCAATGGTCTCATCAACG 3402
      487  TCCATGAGTGTGATGAGCCCATGAATCTCTATTAACCGCGTGAATGACGATCAACAA 546
      3401 GCTCTGGTGTGATGGCCCATTTGAACCTCTAATGCTGTGTGTGACGATGAA 3342
      547  GCCTCTACTAACAAAGAGTGGTGTGATGGAAGATGAGATTCACGCCCCAGAGAA 606
      3341 GACTCTGCAACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3282
      607  GGCACCAAGCTCAACACCAACCGAGTCAATGCAATTTGCTGCCCAACACAGTAAATC 666
      3281 ATCGGTAACTGACACACACTGAAATGTCAGGCTTCCAGGAGTGAACGACGACCA 3222
      667  GGCACAGCTATTATGGAAGTGCAGTATTTCTACTCAATCCGTTGACCTCAGACCTT 726
      3221 GGCCTATTCTATACGTTAAAGTAAAGTAACTACTCTCT---GCACGACCAAGTAAAGAGAT 3165
      727  GCAAGGAGTTGATATTTACAAATCGAAAGACCTCCGAGTGCATATTTCTTACGCT 786
      3164 AAGCTGTTTTGATGTGACCAACTGACTGAACTGCAAAAGTAGTATGTGTATAC 3105
      787  CACCCCGATACATGATGTTTATGTCATGCAATGCAAGCTTTCAGGAGGAGGCAAGAAATC 846
      3104 TACTCTAATGCTTCCGATCTGCGGTAAAGCAATTTGATGAGATGATTAAGAGTAT 3045
      847  ATCATCTCAGGATGGGCAATGGGAAACCTTCCCTTGCCTGACCTAAATGCTTGA 906
      3044 ATCATGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2985
      907  GCACCAATACGAGCTGATGCTGCTGCAAGCTCTGAGTGGGCACTGTGTCCACAC 966
      2984 GCGCTTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2925
      967  CAAAGAGCTGAGTGTGATAGAAATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1026
      2924 CAAATGCTGAAATGATGACAGCAATAGGTTTGTGTGTGTGTGTGTGTGTGTGTGT 2865
      1027 CAAAAGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1086
      2864 CAGGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2805
      1087 CAAAAGATCTTCTCCACCTATTATTCAGAA 1118
      2804 CAGGAATGTTTCTAATATATTATTCCTGAA 2773

```

RESULT 11
us-60-215-161-1295/c

; Sequence 1295, Application us/60215161
; GENERAL INFORMATION:

; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Huesling, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof

```

; FILE REFERENCE: 38-21(51847)A
; CURRENT APPLICATION NUMBER: us/60/215,161
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 1295
; LENGTH: 9082
; ORGANISM: Xenorhabdus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6985)..(9081)
; US-60-215-161-1295

Query Match      26.7%; Score 302.8; DB 75; Length 9082;
Best Local Similarity 58.1%; Pred. No. 3.4e-81;
Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

      127  CCCCAGTACTATCTTACGACAGGAGGACCACTGCTGGTGGGGGAATCAGGCTC 186
      3755  CCAACATTACTGTTGTGGCAACGGGTGGCAAACTGCTGGTGTGTGTGTGTGTGTGT 3696
      187  AAGAGTACTCTGCTGTGAGCAGTCAACCTGTGATAAGCTTCTTGACGCGCTCC 246
      3695  AATCCACTTACCTACGCGGTCAGTGGTGTGATTCATTCGTAATGCTTCCGCT 3636
      247  ATCAAGCACTAGCCACCATCAAGGGTGAACAGATCTCAACATTTGCTCCCAAGATG 306
      3635  ATCAAAACATTGTAATCTGAAAGTGAAGAGGTTGTGTACATCGCTTCCAGACATG 3576
      307  ACGGTGAGTGTGGCTTAACCTAGCCAAAGGCTGTCAATGAGCTCTGCCCAAAAGAG 366
      3575  AACGATCAGTGTGGCTGACCTGTGGCAAAAAATCAATG-----CAGATTGATATA 3522
      367  ACCGAAGCCGTGATCATCAACCATGAACTGACACCATGGAAGAGAACCGCTTCTCTC 426
      3521  ACTGACGGTTTTGTGATCACCCACGCTACAGATACAGGAAGAAAGCTATTCTCTC 3462
      427  AACCTCAGGGTAAAAAGCCAAACCTGCTCTCTTGTGAGGCGCATCGCTCAGGCTCT 486
      3461  GATCTGACACCTCAGTGTCTATTAACCTATGTAAGTGTGGGCAATGCTGTCAATCAAC 3402
      487  TCCATGAGTGTGATGAGCCCATGAATCTCTTAACGCGCTGTAATGCGATCAACAA 546
      3401  GCTCTGGTGTGATGAGCCCATTTGAACCTCTAATGCTGTGTGTGTGTGTGTGTGTGT 3342
      547  GCCTCTACTAACAAAGAGTGGTGTGATGGAAGATGAGATTCACGCCCCAGAGAA 606
      3341  GACTCTGCAACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3282
      607  GGCACCAAGCTCAACACCAACCGAGTCAATGCAATTTGCTTGCACCAAGAGTAAATC 666
      3281  ATCGGTAACTGACCAACCTGAAATGTCAGGCTTTCAGGAGTGAACGACGACGCCACAA 3222
      667  GGCACAGTCTATTATGCAAAAGTGCAGTATTTCTACTCAATTCGCTTGACCTCAGAC 726
      3221  GGCCTATTCTATACGTTAAAGTAAAGTAACTACTCTCT---GCAGCACTGAAAGAGAT 3165
      727  GCAAGTGTGATGATATTAAGCAAAATCGAAGAACTCCCAAGAGTGTATTTCTTACG 786
      3164  AAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3105
      787  CACCCGATGATCTGATGTTTATGTCATGCAAGCTTTCAGGCGCTTGCAGGACGCAAG 846
      3104  TACTCTAATGCTTCCGATCTGCGGTAAAGCAATTTGTGAGAGATATTTAAGGATAT 3045
      847  ATCATCTCAGGATGGGCAATGGGAAACCTTCCCTTGTGACTCAAAATGCTTGA 906
      3044  ATCATGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2985
      907  GCACCAATACGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 966
      2984  GCGCTTAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2925

```


QY 187 AAGATGCTACTCTGCTGGAGCAGTACCGTTGATAGCTTCTTGCAGCCGCTCCGCC 246
 Db 5388 AATCCCGCTACACTGCGGGGATCGAGTGGTTCATCTGTAATGCTGTTCTGCT 5447
 QY 247 ATCAAGACCTAGCCACCATCAAGGATGAACATGCAACATTTGGCTCCAGAGATG 306
 Db 5448 ATCAAAAACATTTGCTAATCTGAAGATGAGCAGTTGTTACCAAGGCTTCAGAGATG 5507
 QY 307 ACAGGTAAGGTGCTTAACTAAGCCAGCGGTCTCAATGATGCTCCGCCAAAAGAG 366
 Db 5508 AAGCATGAGGTGCTGCTGACTGCTGGCAAAAAATTCATG-----CAGATTGATATA 5561
 QY 367 ACCGAGCCGCTGATCATCACCATGGAAGTACACCATGAGAGAGCCGCTTCTCTC 426
 Db 5562 ACTGACGGTTTGGTATCCACCGGTACAGATACAGTGAAGAAAGTCTTATTTCTC 5621
 QY 427 AACCTCAGGTGAAGCAAAACCTGCTGCTTGAAGGCGCATGCGCCAGGCTC 486
 Db 5622 GATCTGACACTGCTGCTCATTAACCTATGCTAATGTTGGCGAATGCTCCATCAAG 5681
 QY 487 TCCATGAGTGTGATGGCCCATCATATCTATTAACCGGTGATGAGCATCAACA 546
 Db 5682 GCTGTGGGTGATGAGGCGCATTAACCTCTATATGCTGTTGTTGAGCTCAGATAA 5741
 QY 547 GCTCTACTAACAAGAGATGCTGATGATGATGATGATGATGATGATGATGATGAT 606
 Db 5742 GACTCTCAAAACCTGCTGCTTACTGCTATGATGATGATGATGATGATGATGATGAT 5801
 QY 607 GCGACAGCTCAACACCGCAGTCAATGCTATTTGCTGCGCCACACAGTAAATC 666
 Db 5802 ATCGGTAACTGACGCAACTGAAAGTCAAGCTTTCAGGAGTGAAGCGGCCACAA 5861
 QY 667 GGCACAGTCTATATGCAAAAGTCAAGTATTTCACTCAATCCGTTGACCTCACACCT 726
 Db 5862 GGTCTTATCTAATACGTAAGTAAAGTAACTACTACTCT---GCAGCAGCTGAAGCAAT 5918
 QY 727 GCAAGTATGTTATATAGCAAAATGAGCAACTCCCGAGAGTGCATTTCTTACGCT 786
 Db 5919 AATCTGTTTTTGTGTCAGCAAACTGACTGATGCTCAAAAGTATGTTATGATAC 5978
 QY 787 CACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
 Db 5979 TACTCTAATGCTTCCGATCTGCGGTTAAAGCATTTGTAGAAATGATTAATTAAGTAT 6038
 QY 847 ATCCATCAGGATGAGGCAATGGAACCTTCCCTTGAATCAAAATGCTCTGAAAAA 906
 Db 6039 ATGATGCTGCTGTTGTAAGGTAACATCATATAGTCTATTTGAGATCTCTGTCACAA 6098
 QY 907 GCAGCCAAATGAGCGTATGCTGCTGCAAGCTCTAGAGTGGCAGTGTTCACACCC 966
 Db 6099 GCGGCTAAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6158
 QY 967 CAAGAGGCTGAAGTGAATGAAGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1026
 Db 6159 CAATATGCTGAAGTGAATGAAGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6218
 QY 1027 CAAAAAGCAGATGCTCTTATGTTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1086
 Db 6219 CAGAAAGTGTGTTCTTCTGCAATTTGCTGCTGACTGCACTGCTGCAACAGCAAAAT 6278
 QY 1087 CAAAAGATCTTCTCACCCTATTAATCCAGAA 1118
 Db 6279 CAGGAAATGTTTCTAATATTAATTAATCTCTGAA 6310

RESULT 14

US-60-215-161-1298
 ; Sequence 1298, Application US/60215161
 ; GENERAL INFORMATION:
 ; APPLICANT: Corbin, David R.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Huesling, Joseph E.
 ; APPLICANT: Krasomil-Osterfeld, Karina C.

; APPLICANT: Malvar, Thomas M.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Spilidonov, Sergei
 ; APPLICANT: Hinkle, Gregory J.
 ; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 ; FILE REFERENCE: 38-21(51847)A
 ; CURRENT APPLICATION NUMBER: US/60/215,161
 ; CURRENT FILING DATE: 2000-06-30
 ; NUMBER OF SEQ ID NOS: 8409
 ; SEQ ID NO 1298
 ; LENGTH: 9082
 ; TYPE: DNA
 ; ORGANISM: Xenorhabdus sp.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (5259)..(6302)
 ; US-60-215-161-1298

Query Match 26.7%; Score 302.8; DB 75; Length 9082;
 Best Local Similarity 58.1%; Pred. No. 3.4e-81;
 Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY 127 CCCCAGTACTATCTGAGCAGAGAGGACGACATCGCTGTTGGGGGATCTAGCGTC 186
 Db 5328 CCAACATTAATCTGTTGGAACCGGTTGGAACGATGATGATGATGATGATGATGATGAT 5387
 QY 187 AAGATGCTACTCTGCTGGAGCAGTACCGTTGATGATGATGATGATGATGATGATGAT 246
 Db 5388 AATCCCGCTACACTGCGGGGATCGAGTGGTTCATCTGTAATGCTGTTCTGCT 5447
 QY 247 ATCAAGACCTAGCCACCATCAAGGATGAACATGCAACATTTGGCTCCAGAGATG 306
 Db 5448 ATCAAAAACATTTGCTAATCTGAAGATGAGCAGTTGTTACCAAGGCTTCAGAGATG 5507
 QY 307 ACAGGTAAGGTGCTTAACTAAGCCAGCGGTCTCAATGATGCTCCGCCAAAAGAG 366
 Db 5508 AAGCATGAGGTGCTGCTGACTGCTGGCAAAAAATTCATG-----CAGATTGATATA 5561
 QY 367 ACCGAGCCGCTGATCATCACCATGGAAGTACACCATGAGAGAGCCGCTTCTCTC 426
 Db 5562 ACTGACGGTTTGGTATCCACCGGTACAGATACAGTGAAGAAAGTCTTATTTCTC 5621
 QY 427 AACCTCAGGTGAAGCAAAACCTGCTGCTTGAAGGCGCATGCGCCAGGCTC 486
 Db 5622 GATCTGACACTGCTGCTCATTAACCTATGCTAATGTTGGCGCAATGCTCCATCAAG 5681
 QY 487 TCCATGAGTGTGATGGCCCATCATATCTATTAACCGGTGATGAGCATCAACA 546
 Db 5682 GCTGTGGGTGATGAGGCGCATTAACCTCTATATGCTGTTGTTGAGCTCAGATATA 5741
 QY 547 GCTCTACTAACAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 606
 Db 5742 GACTCTCAAAACCTGCTGCTTACTGCTATGATGATGATGATGATGATGATGATGATGAT 5801
 QY 607 GCGACAAAGTCAACACCGCAGTCAATGCTATTTGCTGCGCCACACAGTAAATC 666
 Db 5802 ATCGGTAACTGACGCAACTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6218
 QY 667 GGCACAGTCTATTAAGGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6278
 Db 5862 GCTCTAATGCTTCCGATCTGCGGTTAAAGCATTTGTAGAAATGATTAATTAAGTAT 6310
 QY 727 GCAAGTGAAGTGAATGAAGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1026
 Db 5919 CAATATGCTGAAGTGAATGAAGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6218
 QY 787 CACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
 Db 5979 TACTCTAATGCTTCCGATCTGCGGTTAAAGCATTTGTAGAAATGATTAATTAAGTAT 6038
 QY 847 ATCCATCAGGATGAGGCAATGGAACCTTCCCTTGAATCAAAATGCTCTGAAAAA 906
 Db 6039 ATGATGCTGCTGTTGTAAGGTAACATCATATAGTCTATTTGAGATCTCTGTCACAA 6098

	Query Match	Best Local Similarity	Matches 1131: Conservative	99.7%; 99.8%; 0;	Score 1129.8;	DB 6;	Length 1133;	Pred. No. 0;	Indels 2;	Gaps 0;
QY	1	ATGGGCAAGCACCACATCATCATCATATATATGACAGCGCGCTGCTGGCGCGGGGACAGCAT	60							
DB	1	ATGGGCAAGCACCACATCATCATCATATATATGACAGCGCGCTGCTGGCGCGGGGACAGCAT	60							
QY	61	ATGGCTAGACATGACTGCTGTGACAGCAAAATGGGTGCGGGATCCAGCGCCTGTTTGTATG	120							
DB	61	ATGGCTAGACATGACTGCTGTGACAGCAAAATGGGTGCGGGATCCAGCGCCTGTTTGTATG	120							
QY	121	GCTTAACCCCAAGTGACATCTTCTAGCCACAGAGAGCACATCGCTGCTGGGGGGAATCT	180							
DB	121	GCTTAACCCCAAGTGACATCTTCTAGCCACAGAGAGCACATCGCTGCTGGGGGGAATCT	180							
QY	181	AGCGTCAAGATAGTACTACTCTGCTGAGACATCAACCTGTGATTAAGCTTCTTGACAGCGTC	240							
DB	181	AGCGTCAAGATAGTACTACTCTGCTGAGACATCAACCTGTGATTAAGCTTCTTGACAGCGTC	240							
QY	241	CCTGCGCATCAACGACCTTATGCCACCATCAAGGCTGATACAGATCTCAAGCATTTGGCTCCAA	300							
DB	241	CCTGCGCATCAACGACCTTATGCCACCATCAAGGCTGATACAGATCTCAAGCATTTGGCTCCAA	300							
QY	301	GAGATGACAGGGGTAAAGGTGTGGCTTAACTATGCCAAGCGCTGTCATGAGACTCTCGGCCCAA	360							
DB	301	GAGATGACAGGGGTAAAGGTGTGGCTTAACTATGCCAAGCGCTGTCATGAGACTCTCGGCCCAA	360							
QY	361	AAAGAGACCGAAGCGGTGATCATCAACCCATGAACTGACACCATGGAAGAGACCGCTTTC	420							
DB	361	AAAGAGACCGAAGCGGTGATCATCAACCCATGGAAGAGAGACCGCTTTC	420							
QY	421	TTCTCTCAACCTCAAGGTGAAAAGCCAAAACCTGTCTGCTTGTATGAGCGGCATCGCTCCA	480							
DB	421	TTCTCTCAACCTCAAGGTGAAAAGCCAAAACCTGTCTGCTTGTATGAGCGGCATCGCTCCA	480							
QY	481	GGCTCTTCCATGAGTGTGATGAGCCCATGATCTCTATTAACGGCGTGAATGTAGCGATC	540							
DB	481	GGCTCTTCCATGAGTGTGATGAGCCCATGATCTCTATTAACGGCGTGAATGTAGCGATC	540							
QY	541	AACAAAGCCTTACTTAACAAAGAGGTGTATTTGTATGAAACGATGAGATTCACGCGCC	600							
DB	541	AACAAAGCCTTACTTAACAAAGAGGTGTATTTGTATGAAACGATGAGATTCACGCGCC	600							
QY	601	AGAAAGAGGACCAAGCTCAACACACACGCAATCATGATTCGTCGCCAACACAGGT	660							
DB	601	AGAAAGAGGACCAAGCTCAACACACACGCAATCATGATTCGTCGCCAACACAGGT	660							
QY	661	AAAATCGGCACAGTCTATTATGCGAAAGTGCAGATTTTCACTCAATCCGTTTCGACTCAC	720							
DB	661	AAAATCGGCACAGTCTATTATGCGAAAGTGCAGATTTTCACTCAATCCGTTTCGACTCAC	720							
QY	721	ACCGTTGCAATGAGTTGATTAATAGCAAAATCGAAGATCCCAAGAGTGTGATATTCCT	780							
DB	721	ACCGTTGCAATGAGTTGATTAATAGCAAAATCGAAGATCCCAAGAGTGTGATATTCCT	780							
QY	781	TACGCTCACCCCGATGATCTGATGTTTATGTCATGACGACCCCTTACAGGACAGACAAA	840							
DB	781	TACGCTCACCCCGATGATCTGATGTTTATGTCATGACGACCCCTTACAGGACAGACAAA	840							
QY	841	GGATTCATTCATGAGAGCATGGGAATGGGAACCCCTTCCCTTGACATCAAAATGCTCTT	900							
DB	841	GGATTCATTCATGAGAGCATGGGAATGGGAACCCCTTCCCTTGACATCAAAATGCTCTT	900							
QY	901	GAAGAAACAGCCCAATAGAGCGTATGCTGCTCGAAGCTCTAGAGTGGGACAGTGGTTC	960							
DB	901	GAAGAAACAGCCCAATAGAGCGTATGCTGCTCGAAGCTCTAGAGTGGGACAGTGGTTC	960							
QY	961	ACCACCCAGAGGCTGAAGTGTGATGATPAGAACTTGGTTTGTGGCTACAGAGATCTC	1020							

Dd	961	AOCACCAAGAGGCTGAGTGGATGATATAAAGAAACTTGTTGTGTGGCTACAGAGAGTCTC	1021
Qy	1021	AACCTCAA AAAAGCCAGAGTGGCTTTATATGTTAGCGCTTCACC AAATAGTAGTAGAGAG	1080
Dd	1021	AACCTCAA AAAAGCCAGAGTGGCTTTATATGTTAGCGCTTCACC AAATAGTAGTAGAGAG	1080
Qy	1081	GCGATCCAAAAGATCTTCCACCACTATTAATCCAGAAGAAAGGAATCTCTCAC	1133
Dd	1081	GCGATCCAAAAGATCTTCTCCACCTATTATCCAGAAAGGAATCTCTCAC	1133
 RESULT 2 US-09-897-516A-1293/c ; Sequence 1293, Application US/09897516A ; GENERAL INFORMATION: ; APPLICANT: Corbin, David R. ; APPLICANT: Goldman, Barry S. ; APPLICANT: Hinkle, Gregory J. ; APPLICANT: Hueising, Joseph E. ; APPLICANT: Malvar, Thomas M. ; APPLICANT: Krasomil-Osterfeld, Karina C. ; APPLICANT: Slater, Steven C. ; APPLICANT: Spiridonov, Sergei ; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof ; FILE REFERENCE: 38-21(51847)B ; CURRENT APPLICATION NUMBER: US/09/897,516A ; PRIOR FILING DATE: 2001-06-29 ; PRIOR APPLICATION NUMBER: US 60/215,161 ; PRIOR FILING DATE: 2000-06-30 ; NUMBER OF SEQ ID NOS: 8415 ; SEQ ID NO 1293 ; LENGTH: 9082 ; TYPE: DNA ; ORGANISM: Xenorhabdus sp. ; FEATURE: ; NAME/KEY: CDS ; LOCATION: (3968)..(5734) ; OTHER INFORMATION: US-09-897-516A-1293			
 Query Match 26.7%; Score 302.8; DB 5; Length 9082; Best Local Similarity 58.1%; Pred. No. 2.7e-91; Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;			
Qy	127	CCCCAGAGCATCTATCTCTGAGCCACAGAGGACGACACCTCGTGTGGGGGGAATCTAGCTC	186
Dd	3755	CCAAACATTACTGTCTTGCGCAACGGGTGGCAACAATGCTGGTGGTGAATCTGCACAT	3696
Qy	187	AAGAGTACTCTACTCTGCTGGAGACAGTCAACCTTGATTAAGCTTCTTGACGCCCTCCCTCC	246
Dd	3695	AAATCCACTACACTGACGGGGTCAAGTGGCTGTGTTGTTATGTTGAATGCTGTCTCTCT	3636
Qy	247	ATCAACGACCTTAGCCACCATCAAGGGTGAACAGATCTCAGCATTTGGCTCCCAAGAGATG	306
Dd	3635	ATCAAAAAACATTTGCTAAATCTGAAGAGGTAGCAGAGTTGTTACATCGCTTCAGAGACATG	3576
Qy	307	ACGGGTAAAGTGGCTTAAACTAGGCCAAGGGTGCATGAGCTCTCGGCCCAAAAAAGAG	366
Dd	3575	AACGATCAGGCTGTGGCTACTCTGTGCGCAAAAAAATCAATG-----CAGATTGTGATGAA	3522
Qy	367	ACCGAAGCCGTGATCATCACCATGGAAGTGAACATGACACATGAGAAAGACCGCTTCTTCCTC	426
Dd	3521	ACTGACGGTTTTTGTCATCACCACAGGGTACAGATACCATGGAAGAAACGTCTATTTCCTC	3462
Qy	427	AACTCAGCGTGAAGAACCAAAACCTCTGTGCTTGTAGGCGCCATCGTCCAGGCTCT	486
Dd	3461	GATCTGACCACTCAGTGTCTATTAACCTATGTAATGTTGGCGCAATCGTCCATCAACG	3402
Qy	487	TTCATGATGCTGATGGCCCATGATGATCTATTAAGCCCGGAATGTGCGATCAACAA	546
Dd	3401	GCTCTGGTGTGATGAGCCATTGAACCTCTATATATGCTGTGTTGTAAGCTTCAGATGAA	3342

[illegible]

RESULT 4
US-09-897-516A-1295/c
; Sequence 1295, Application US/09897516A

GENERAL INFORMATION:
 APPLICANT: Corbin, David R.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Huesing, Joseph E.
 APPLICANT: Malvar, Thomas M.
 APPLICANT: Krasomil-Osterfeld, Karina C.
 APPLICANT: Slater, Steven C.
 TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 FILE REFERENCE: 38-21(51847)B
 CURRENT APPLICATION NUMBER: US/09/897,516A
 PRIOR FILING DATE: 2001-06-29
 PRIOR APPLICATION NUMBER: US 60/215,161
 NUMBER OF SEQ ID NOS: 8415
 SEQ ID NO 1295
 LENGTH: 9082
 TYPE: DNA
 ORGANISM: Xenorhabdus sp.
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (6985)..(9081)
 OTHER INFORMATION:
 US-09-897-516A-1295

Query Match 26.7%; Score 302.8; DB 5; Length 9082;
 Best Local Similarity 58.1%; Pred. No. 2.7e-91;
 Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

127 CCCAAGTACATCTCTAGCCAGAGGACCATCGTGGGGAATCTAGCGTC 186
 3755 CCACAACTTACTCTCTGGCAAGGGGTGACAACTGCTGGTGAATCTGCACT 3696
 187 AAGAGTACCTCTCTGGGAGGACGACCTGATTAAGCTCTGCAAGCTGCTGCT 246
 3695 AATTCAGCTACCTCTGGGAGGACGACCTGATTAAGCTCTGCAAGCTGCTGCT 3636
 247 ATCAACGACCTACGACCATCAAGGATGAGTCAAGCTGATTAAGCTCTGCAAG 306
 3635 ATCAACGACCTACGACCATCAAGGATGAGTCAAGCTGATTAAGCTCTGCAAG 3576
 307 ACGGGTACCTCTGATTAAGCTCAAGGATGAGTCAAGCTGATTAAGCTCTGCA 366
 3575 AAGCATCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3522
 367 ACCGAGCCGTGATCATCACCATGAGTGAACATGAGTGAAGAGAGAGAGAGAG 426
 3521 ACTGACGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3462
 427 AACCTCAGGTAAGGAAAGGCAAAACCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
 3461 GATCTGACCACTCAGTGTCAATTAACCTGATTAAGCTGATTAAGCTGATTA 3402
 487 TCCATGAGTCTGATGAGGACCATGATTAAGCTGATTAAGCTGATTAAGCTG 546
 3401 GCTGTGGGTGCTGATGAGGACCATGATTAAGCTGATTAAGCTGATTAAGCT 3342
 547 GCTGTGATTAAGGAGAGGATGATTAAGCTGATTAAGCTGATTAAGCTGATTA 606
 3341 GACTCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3282
 607 GCGACCAAGCTCAACACCCGAGTCAATGATTAAGCTGATTAAGCTGATTA 666
 3281 ATCGGTAACTAGGACCAACCTGATTAAGCTGATTAAGCTGATTAAGCTGAT 3222
 667 GCGACAGTCTATTAAGGCAAGTGAATTAAGCTGATTAAGCTGATTAAGCTG 726
 3221 GCGTCTATTAAGGCAAGTGAATTAAGCTGATTAAGCTGATTAAGCTGATTA 3165
 727 GCAAGTGAATTAAGGCAAGTGAATTAAGCTGATTAAGCTGATTAAGCTGAT 786
 3164 AAGGCTGATTAAGGCAAGTGAATTAAGCTGATTAAGCTGATTAAGCTGATTA 3105

787 CACCCGATGATCTATGTTTATGATGACGCTTCAGGAGGACCAAGAGATC 846
 3104 TACTCTATGCTCTTCGATCTGCGGTTTAAAGCTTTGTAGAGATGATTAAGGAT 3045
 847 ATCCATGACGATGAGGCAATGAGGACCTTCCTTACCTCAAAATCTCTGAAAA 906
 3044 ATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2985
 907 GCGACCAATCAGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
 2984 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2925
 967 CAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026
 2924 CAAATGCTGATTAAGGCAATGAGGACCTTCCTTACCTCAAAATCTCTGAAAA 2865
 1027 CAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1086
 2864 CAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2805
 1087 CAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1118
 2804 CAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2773

Result 5
 US-09-897-516A-1296
 Sequence 1296, Application US/09897516A

GENERAL INFORMATION:
 APPLICANT: Corbin, David R.
 APPLICANT: Goldman, Barry S.
 APPLICANT: Hinkle, Gregory J.
 APPLICANT: Huesing, Joseph E.
 APPLICANT: Malvar, Thomas M.
 APPLICANT: Krasomil-Osterfeld, Karina C.
 APPLICANT: Slater, Steven C.
 TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 FILE REFERENCE: 38-21(51847)B
 CURRENT APPLICATION NUMBER: US/09/897,516A
 PRIOR FILING DATE: 2001-06-29
 PRIOR APPLICATION NUMBER: US 60/215,161
 NUMBER OF SEQ ID NOS: 8415
 SEQ ID NO 1296
 LENGTH: 9082
 TYPE: DNA
 ORGANISM: Xenorhabdus sp.
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (8058)..(9080)
 OTHER INFORMATION:
 US-09-897-516A-1296

Query Match 26.7%; Score 302.8; DB 5; Length 9082;
 Best Local Similarity 58.1%; Pred. No. 2.7e-91;
 Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

127 CCCAAGTACATCTCTAGCCAGAGGACCATCGTGGGGAATCTAGCGTC 186
 5328 CCACAACTTACTCTCTGGCAAGGGGTGACAACTGCTGGTGAATCTGCACT 5387
 187 AAGAGTACCTCTCTGGGAGGACGACCTGATTAAGCTCTGCAAGCTGCTGCT 5447
 5388 AATTCAGCTACCTCTGGGAGGACGACCTGATTAAGCTCTGCAAGCTGCTGCT 306
 247 ATCAACGACCTACGACCATCAAGGATGAGTCAAGCTGATTAAGCTCTGCAAG 306
 5448 ATCAACGACCTACGACCATCAAGGATGAGTCAAGCTGATTAAGCTCTGCAAG 5507
 307 ACGGTAAGTGTGCTTAACTAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTG 366

Db 6159 CAAATGCTGAAGTAAATGACGCAAAATGCGTTTGGTCATCAGAACGCGTGAACCCG 6218
 QY 1027 CAAAAGCCAGAGCTCTCTTATGTTAGCCCTCACCACAAAGTAGATGAGAGGCGATC 1086
 Db 6219 CAGAAAGCTCGGTTCTCTGCAATTTGGCTGACCTCAGACGCTGACAGACCAAAATTT 6278
 QY 1087 CAAAGATCTCTCCACCTATTAATCCAGAA 1118
 Db 6279 CAGGAATGTTTCTTAATTAATTCCTGAA 6310

RESULT 7

US-09-897-516A-1298
 : Sequence 1298, Application US/09897516A
 : GENERAL INFORMATION:
 : APPLICANT: Corbin, David R.
 : APPLICANT: Goldman, Barry S.
 : APPLICANT: Hinkle, Gregory J.
 : APPLICANT: Huesing, Joseph E.
 : APPLICANT: Malvar, Thomas M.
 : APPLICANT: Krasomil-Osterfeld, Karina C.
 : APPLICANT: Slater, Steven C.
 : TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
 : FILE REFERENCE: 38-21(51847)B
 : CURRENT APPLICATION NUMBER: US/09/897,516A
 : PRIOR FILING DATE: 2001-06-29
 : PRIOR APPLICATION NUMBER: US 60/215,161
 : NUMBER OF SEQ ID NOS: 8415
 : SEQ ID NO 1298
 : LENGTH: 9082
 : TYPE: DNA
 : ORGANISM: Xenorhabdus sp.
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (5329)..(6302)
 : OTHER INFORMATION:
 : US-09-897-516A-1298

Query Match 26.7%; Score 302.8; DB 5; Length 9082;
 Best Local Similarity 58.1%; Pred. No. 2,7e-91;

Matches 576; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY 127 CCCCAGTACATCTTCTAGCCACAGAGCACCACCTCGTGGGGGATCTAGAGTC 186
 Db 5328 CCAACATCTACTGTTCTGCAAGGGGAGCACAATCGCTGCTGTAATCTGCACT 5387
 QY 187 AAGAGTAGTCTCTGCTGAGCAGTCACCCGTTGATAGCTTTGCGAGCCCTGCC 246
 Db 5388 AATTCAGCTACACTCGGGTCAAGTCGGTGTGATTCATTCGTAATGCTTCTGCT 5447
 QY 247 ATCAAGACCTAGCCACATCAAGGGTGAAGATCAATCAATGCTGGCCCAAGAGATG 306
 Db 5448 ATCAAAATATCTCTATCTGAAGAGTGAACAGCTGTTGATGCAATGCTTCCAGCAT 5507
 QY 307 ACGGGTAGGTGCTTAACTAGCCAAAGCGTGCATGATGAGCTCCCGCCAAAAGAG 386
 Db 5508 AACGATCAGGTCTGCTGACTCGGCAAAAATCAATG-----CAGATTGTGATTA 5561
 QY 367 ACCGAAGCGGTGATCATCAACCATGAGTACACCATGAGAGAGACCGCTTCTCTC 426
 Db 5562 ACTGACGGTTTGTCTATCAACCAAGGTACAGATACCATGGAAGAACTCTTATTTCTC 5621
 QY 427 AACCTCAGCGTAAAAAGCCAAACCTGTCTGCTTGTAGGGCCATGGCTCCAGGCTCT 486
 Db 5622 GATCTGACCATCTAGTGTATTAACCTATCTGTAATGTTGGCCCATGGCTCATTAAG 5681
 QY 487 TCCATGAGGTCTGATGAGCCCAATCTCTTAAAGCCGCTGAATGATGACGATCAACA 546
 Db 5682 GCTCTGGGCTGATGAGCCCAATCTCTTAAATGCTGTGTGTACCGTCAGATTA 5741
 QY 547 GCCTTACTAACAAGAGTGTGATTTGATGAAGATGATTCACCGCCCGCAGAGAA 606

Db 5742 GACTCTGCAAAACCGTGTCTTTTACTTGTATGAAGATTCGTAAATCATGAGACGTGAC 5801
 QY 607 GCGACCAAGCTCAACACCAACCCAGTCAATGATTTGCTGGCCCAACAGGTAAATC 666
 Db 5802 ATCGTAACCTGAGCAACCTGAATGAGGCTTCCAGAGCATGACGACGACCCACAA 5861
 QY 667 GGCACAGTCTATATGCGCAAGTCCAGATTTTCACTCAATCCGTTGACCTCACAACCT 726
 Db 5862 GCTTCATCTAATGATGAAGTAAAGTAACTACTCT---GCAGCACAGATGAACAGAT 5918
 QY 727 GCAGTGTATGATTAATGCAAAATGCAAGATCCCGAGTGTGATTTCTTACGT 786
 Db 5919 AAGCTGTTTGTGATGACGAACTGACTGACTGCCAAAGATGATTTGTATTAAC 5978
 QY 787 CACCCGATGATCTGATGTTTGTAGTCATGAGCCCTTCAGGACGAGCCAAAGATC 846
 Db 5979 TACTCTAATGCTCCGATCTGCGGTAAAGCATTTGATGAAGATGATTAAGGTATT 6038
 QY 847 ATCATGAGCATGGGCAATGGGAACCTTCCCTTGACTCAAAATGCTTTGAAAA 906
 Db 6039 ATCATGCTGCTGTTGTTGTAACGTAATCTATTAAGTCTATTCGACTCTCTGCACAA 6098
 QY 907 GCAGCCAAATCAGCGGTAGTGTGCTGCAAGCTTGAAGTGGGAGGTTCACACCC 966
 Db 6099 GCGGCTAAAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6158
 QY 967 CAAGGCTGAAGTATGATTAAGAACTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1026
 Db 6159 CAAATGCTGAAGTAAAGACGCAAAATGCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6218
 QY 1027 CAAAAGCCAGAGTCTCTTATGTTAGCCCTCACCACAAATAGATGATGAGAGGCGATC 1086
 Db 6219 CAGAAAGCTGTCTTCTTCTCAATTTGGCTGACTGACTGACTGACTGACTGACTGACTG 1118
 QY 1087 CAAAGATCTCTCACCCTATTAATCCAGAA 6278
 Db 6279 CAGGAATGTTTCTTAATTAATTCCTGAA 6310

RESULT 8

US-10-329-670-1
 : Sequence 1, Application US/10329670

: GENERAL INFORMATION:

: APPLICANT: Fleischmann et al.

: TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, F

: FILE REFERENCE: P186P1

: CURRENT APPLICATION NUMBER: US/10/329,670

: PRIOR APPLICATION NUMBER: US 09/643,990

: PRIOR FILING DATE: 2000-08-23

: PRIOR APPLICATION NUMBER: US 08/487,429

: PRIOR FILING DATE: 1995-06-07

: PRIOR APPLICATION NUMBER: US 08/426,787

: NUMBER OF SEQ ID NOS: 1

: SOFTWARE: Patent version 3.1

: SEQ ID NO 1

: LENGTH: 1830121

: TYPE: DNA

: ORGANISM: Haemophilus influenzae

: FEATURE:

: NAME/KEY: misc_feature

: LOCATION: (4747)..(4747)

: OTHER INFORMATION: n equals a, t, g or c

: NAME/KEY: misc_feature

: LOCATION: (9921)..(9921)

: OTHER INFORMATION: n equals a, t, g or c

: NAME/KEY: misc_feature

: LOCATION: (10150)..(10150)

OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40810)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (13360)..(13360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (13910)..(13910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:


```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (b) LOCATION 1...966
; SEQUENCE DESCRIPTION: SEQ ID NO: 791:
US-10-617-320-791

Query Match          7.48; Score 84.4; DB 6; Length 966;
Best Local Similarity 51.9%; Pred. No. 8.6e-18;
Matches 190; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 369 CGAAGCGGTATGATGACCCATGAGACGACACCATGGAAGAGACGGCTTCTCTCA 428
DB 231 CGATGGAGTGTGATGACACGACGACGACGACGACGACGACGACGACGACGACGAC 290
QY 429 CCTCAGCGTGAAGGACGACGACGACGACGACGACGACGACGACGACGACGAC 488
DB 291 TACCATGGAAGTCCCATATGCTATGCTTAACAGAGGACGACGACGACGACGAC 350
QY 489 CATGAGTGTGATGACGACGACGACGACGACGACGACGACGACGACGACGAC 548
DB 351 GCTCGGAGTGTGATGACGACGACGACGACGACGACGACGACGACGACGACGAC 410
QY 549 CTGCTAACAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 608
DB 411 GCGCTGACACAAAGAGATTTTGTGCTATGACGATGAATCCACGCTGACGACGAT 470
QY 609 GACCAAGCTCAACACGACGACGACGACGACGACGACGACGACGACGACGAC 668
DB 471 CACCAAAACATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 530
QY 669 CACAGCTATTTATGGAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 728
DB 531 TCTCATATGAAACAGAAATCTCTACTTCAAAACAGCTGACCTGCTGCTTGA 590
QY 729 AAGTGA 734
DB 591 CCTTGA 596

RESULT 12
US-10-406-676-3
; Sequence 3, Application US/10406676
; GENERAL INFORMATION:
; APPLICANT: Structural Genomics, Inc.
; APPLICANT: Antonysamy, Stephen
; APPLICANT: Fell, Ingeborg
; APPLICANT: Buchanan, Sean
; APPLICANT: Post, Kai W.
; APPLICANT: Liu, Yi
; APPLICANT: Lober, David
; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF PAK4KD KINASE
; TITLE OF INVENTION: PAK4KD
; FILE REFERENCE: 524982002300
; CURRENT APPLICATION NUMBER: US/10/406,676
; CURRENT FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 60/371,018
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/430,567
; PRIOR FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 82

```

```

; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence after ligation into vector
US-10-406-676-3

Query Match          5.98; Score 67.4; DB 6; Length 82;
Best Local Similarity 98.6%; Pred. No. 1.5e-12;
Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGACGACGACCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 60
DB 14 ATGGGACGACGACCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 73
QY 61 ATGGCTAGC 69
DB 74 ATGGCTAGC 82

RESULT 13
PCT-US03-06661A-9
; Sequence 9, Application PC/TUS0306661A
; GENERAL INFORMATION:
; APPLICANT: Chou, Szu-Yi
; TITLE OF INVENTION: Method of Producing Antigens
; FILE REFERENCE: SAMG/0002 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/06661A
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 60/361,166
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/363,445
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: 10/231,114
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,470
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,063
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,213
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/231,298
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent version 3.1
; SEQ ID NO 9
; LENGTH: 1067
; TYPE: DNA
; ORGANISM: Streptomyces mobaraensis ATCC 29032
PCT-US03-06661A-9

Query Match          5.98; Score 67.4; DB 1; Length 1067;
Best Local Similarity 98.6%; Pred. No. 5.2e-12;
Matches 68; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGACGACGACCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 60
DB 1 ATGGGACGACGACCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 60
QY 61 ATGGCTAGC 69
DB 61 ATGGCTAGC 69

RESULT 14
PCT-US03-06661A-11
; Sequence 11, Application PC/TUS0306661A
; GENERAL INFORMATION:
; APPLICANT: Chou, Szu-Yi
; TITLE OF INVENTION: Method of Producing Antigens
; FILE REFERENCE: SAMG/0002 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/06661A
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 60/361,166

```

Query Match	5.9%	Score 67.4	DB 1	Length 1074
Best Local Similarity	98.6%	Pred. No. 5.2e-12		
Matches 68; Conservative	0	Mismatches 1	Indels 0	Gaps 0

RESULT 15
US-10-423-156-60
; Sequence 60, Application US/10423156
; GENERAL INFORMATION

```

1  APPLICANT: Lin, Hsin-Yu
2  APPLICANT: Hwang, Ching-Long
3  TITLE OF INVENTION: ANTIGENIC FRAGMENT OF HUMAN
4  TITLE OF INVENTION: T-LYMPHROTROPIC VIRUS
5  FILE REFERENCE: 05204-020001
6  CURRENT APPLICATION NUMBER: US/10/423,156
7  CURRENT FILING DATE: 2003-04-25
8  PRIOR APPLICATION NUMBER: TW 91135980
9  PRIOR FILING DATE: 2002-12-12
10 NUMBER OF SEQ ID NOS: 60
11 SOFTWARE: FastSeq for Windows Version 4.0
12 SEQ ID NO 60
13 LENGTH: 759
14 TYPE: DNA
15 ORGANISM: Artificial Sequence
16 FEATURE:
17 OTHER INFORMATION: Synthetically generated oligonucleotide
18
19 SEQ-10-023-156-60

```

Query Match	5.8%	Score 65.2	DB 6	Length 759
Best Local Similarity	95.7%	Pred. No. 2.4e-11		
Matches 67	Conservative	0	Mismatches 3	Indels 0
				Gaps 0

QY 1 ATGGGAGAGGCATCATCATCATCATCATATGAGACAGGGGCGTGGTGGCGGGGGAGCCAT 60
 |||||
 364 ATGGGAGAGGCATCATCATCATCATCATATGAGACAGGGGCGTGGTGGCGGGGGAGCCAT 423
 Db
 QY 61 ATGGCTAGCA 70
 |||||
 424 ATGGCTGCCA 433
 Db

Search completed: September 4, 2003, 19:14:51
Job time : 217 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2003, 16:08:40 ; Search time 4197 seconds

(without alignments)
11043.747 Million cell updates/sec

Title: US-09-937-982-3

Perfect score: 1133

Sequence: 1 atgggcagcagccatcatca.....aagaagggatctcttcac 1133

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl : *

1: gb_ba : *

2: gb_hlg : *

3: gb_in : *

4: gb_om : *

5: gb_ov : *

6: gb_pat : *

7: gb_ph : *

8: gb_pl : *

9: gb_pr : *

10: gb_ro : *

11: gb_sts : *

12: gb_sy : *

13: gb_un : *

14: gb_vi : *

15: em_ba : *

16: em_fun : *

17: em_hum : *

18: em_in : *

19: em_mu : *

20: em_om : *

21: em_or : *

22: em_ov : *

23: em_pat : *

24: em_ph : *

25: em_pl : *

26: em_ro : *

27: em_sts : *

28: em_un : *

29: em_vi : *

30: em_hlg_hum : *

31: em_hlg_inv : *

32: em_hlg_other : *

33: em_hlg_mus : *

34: em_hlg_pin : *

35: em_hlg_rod : *

36: em_hlg_mam : *

37: em_hlg_vrt : *

38: em_sy : *

39: em_hgo_hum : *

40: em_hgo_mus : *

41: em_hgo_other : *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1129.8	99.7	1133	6	AR157944
2	1129.8	99.7	1133	6	AX464427
3	1027.8	90.7	1035	6	WSANSAGEN
4	1027.8	90.7	1035	6	WSDCUNSA
5	368.8	32.6	3606	1	WSAJ2933
6	323.2	28.5	303450	1	AP005085
7	300	26.5	10893	1	AE015308
8	300	26.5	292906	1	AE016988
9	296.8	26.2	10749	1	AE005526
10	296.8	26.2	266658	1	AE002563
11	296.8	26.2	303121	1	AE016766
12	296.4	26.2	10173	1	U32758
13	296.4	26.2	110000	6	AR274513.07
14	296.4	26.2	110000	6	AR274513.08
15	295.2	26.1	1530	1	ECOLASNTT
16	295.2	26.1	1643	1	ECOANSBA
17	295.2	26.1	10448	1	AE000378
18	295.2	26.1	141744	1	ECU28377
19	294	25.9	1848	6	I73512
20	294	25.9	1848	12	U06943
21	283.6	24.9	10769	1	AE013881
22	282.6	24.9	210050	1	AJ414148
23	276	24.4	22204	1	AE008842
24	276	24.4	230050	1	AL627277
25	276	24.4	301574	1	AE016844
26	273.2	24.1	305961	1	AE016937
27	270.4	23.9	10094	1	AE001498
28	259.2	22.9	11196	6	BD092628
29	259.2	22.9	10923	1	AE000585
30	230.4	20.3	304500	1	AP005953
31	224.8	19.8	2450	1	ECSC
32	223.2	19.7	2837	6	A14577
33	223.2	19.7	2837	6	E01113
34	221	19.5	35000	1	AE000617
35	221	19.5	221160	1	BSUB0002
36	215.2	19.0	1399	1	ECASN
37	210.8	18.6	1017	6	AR177619
38	209.2	18.5	1709	1	AF056495
39	207.6	18.3	314150	1	CJ11168X1
40	205.2	18.1	308015	1	AE016783
41	203.2	17.9	110000	2	AE020884.1
42	200.4	17.7	9937	1	AE004563
43	195.4	17.2	2353	1	PFL238710
44	178.8	15.8	20516	1	AE008866
45	128.8	11.4	1337	8	SPLASPA

ALIGNMENTS

RESULT 1

LOCUS AR157944 1133 bp DNA

DEFINITION Sequence 3 from patent US 6251388.

ACCESSION AR157944

VERSION AR157944.1 GI:16219888

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 1133)

AUTHORS Durdan,D.L.

TITLE Utilization of Wolinella succinogenes asparaginase to treat diseases associated with asparagine dependence

JOURNAL Patent: US 6251388-A 3 26-JUN-2001;

Pred. No. is the number of results predicted by chance to have a

FEATURES Location/Qualifiers
 source 1. 1133 /organism="unknown"
 BASE COUNT 315 a 305 c 267 g 246 t
 ORIGIN

Query Match 99.7%; Score 1129.8; DB 6; Length 1133;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1131: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGGGCGACGAGCCATCATCATCATCATAGACAGCGCGCTGGTCCGGCGGAGCCAT 60
DB 1 ATGGGCGACGAGCCATCATCATCATCATAGACAGCGCGCTGGTCCGGCGGAGCCAT 60
QY 61 ATGGCTGACATGATGCTGGGACAGCAAAATGGGTGCGGATCCAGGCGCTCTGTTTGATG 120
DB 61 ATGGCTGACATGATGCTGGGACAGCAAAATGGGTGCGGATCCAGGCGCTCTGTTTGATG 120
QY 121 GCTAAACCCCAAGTACTATCTATCTAGCCACAGAGAGGACACCATCGTGGTGGGGAATCT 180
DB 121 GCTAAACCCCAAGTACTATCTATCTAGCCACAGAGAGGACACCATCGTGGTGGGGAATCT 180
QY 181 AGCGTCAAGAGTACTACTCTGCTGAGACAGTCAACCGTTGATAGCTTTGACAGCGTTC 240
DB 181 AGCGTCAAGAGTACTACTCTGCTGAGACAGTCAACCGTTGATAGCTTTGACAGCGTTC 240
QY 241 CCTGCGATCAACGACCTAGCCACCATCAAGGGGTGAAGAGATCTCAAGCATTTGGCTCCCA 300
DB 241 CCTGCGATCAACGACCTAGCCACCATCAAGGGGTGAAGAGATCTCAAGCATTTGGCTCCCA 300
QY 301 GAGATGACGGGTAAAGGTGGCTTAACTAGCCCAAGCGGTGCAATGAGCTCTCGGCCAA 360
DB 301 GAGATGACGGGTAAAGGTGGCTTAACTAGCCCAAGCGGTGCAATGAGCTCTCGGCCAA 360
QY 361 AAAGAGACCGGAAGCGGTGATCATCAACCATGAGACATGAGACATGAGAGACCGGTTTC 420
DB 361 AAAGAGACCGGAAGCGGTGATCATCAACCATGAGACATGAGACATGAGAGACCGGTTTC 420
QY 421 TTCTCTCAACCTCAGCGGTGAAAAAGCCAAACCTGTCTGCTGTAGAGCGCCATGCTGCA 480
DB 421 TTCTCTCAACCTCAGCGGTGAAAAAGCCAAACCTGTCTGCTGTAGAGCGCCATGCTGCA 480
QY 481 GGCCTCTTCATGATGCTGATGAGGCGCCATGATCTCTATAGAGCGCGTGAATGAGGATC 540
DB 481 GGCCTCTTCATGATGCTGATGAGGCGCCATGATCTCTATAGAGCGCGTGAATGAGGATC 540
QY 541 AACCAAGCGCTACTACAAAGAGAGGATGATGATGATGATGATGATGATGATGATGATG 600
DB 541 AACCAAGCGCTACTACAAAGAGAGGATGATGATGATGATGATGATGATGATGATGATG 600
QY 601 AGAGAAAGCGACCAAGCTCAACACACACGAGTCAATGATGCTTCCGCCCAACAGAGT 660
DB 601 AGAGAAAGCGACCAAGCTCAACACACACGAGTCAATGATGCTTCCGCCCAACAGAGT 660
QY 661 AAAATCGGCACAGTCTATTTATGCAAAAGTCAAGTATTTACTCAATCCGTTGACCTCAC 720
DB 661 AAAATCGGCACAGTCTATTTATGCAAAAGTCAAGTATTTACTCAATCCGTTGACCTCAC 720
QY 721 ACCCTGCAAGTGAATGATATGCAAAATGCAAGAACTCCCGAGAGTCAATTTCTT 780
DB 721 ACCCTGCAAGTGAATGATATGCAAAATGCAAGAACTCCCGAGAGTCAATTTCTT 780
QY 781 TACGCTCAACCCGATGATAGTATGTTTATGTCATATGACGACCGCTTCAGCGAGAGCCAAA 840
DB 781 TACGCTCAACCCGATGATAGTATGTTTATGTCATATGACGACCGCTTCAGCGAGAGCCAAA 840
QY 841 GGAATCATTCATGACGAGGATGAGGCAATGAGAACCTTTCCCTTTGATCAAAATGCTCTT 900
DB 841 GGAATCATTCATGACGAGGATGAGGCAATGAGAACCTTTCCCTTTGATCAAAATGCTCTT 900
QY 901 GAAAAAGACGCAAAATCAAGGCTAGTCTGCTGCAAGCTCTAGAGTGGGAGTGGTTCC 960
DB 901 GAAAAAGACGCAAAATCAAGGCTAGTCTGCTGCAAGCTCTAGAGTGGGAGTGGTTCC 960

```

QY 961 ACCACCAAGAGGCTGAAGTATGATATGAAGAACTGGTTTGTGCTACAGAGATCTC 1020
 DB 961 ACCACCAAGAGGCTGAAGTATGATATGAAGAACTGGTTTGTGCTACAGAGATCTC 1020

QY 1021 AACCTCCAAAAGCCAGAGTCTTCTTATGTTTACCCCTCACCAAACTAGTATGAGAG 1080
 DB 1021 AACCTCCAAAAGCCAGAGTCTTCTTATGTTTACCCCTCACCAAACTAGTATGAGAG 1080

QY 1081 GCGATCCAAAAGATCTTCCACCTATTAATCCAAAGAGGAAATCTCTCAC 1133
 DB 1081 GCGATCCAAAAGATCTTCCACCTATTAATCCAAAGAGGAAATCTCTCAC 1133

RESULT 2
 AX464427 1133 bp DNA linear PAT 16-JUL-2002
 LOCUS
 DEFINITION Sequence 3 from Patent EPI219706.
 AX464427
 VERSION AX464427.1 GI:21899232
 KEYWORDS
 SOURCE Wolinella succinogenes
 ORGANISM Wolinella succinogenes
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 Helicobacteriaceae; Wolinella.

REFERENCE
 AUTHORS Durdan, D.L.
 TITLE Production processes for recombinant wolinella succinogenes
 JOURNAL asparaginase
 PATENT: EP 1219706-A 3 03-JUL-2002;
 CHILDRENS HOSPITAL OF LOS ANGELES (US)
 FEATURES
 source 1. 1133
 /organism="Wolinella succinogenes"
 /mol_type="genomic DNA"
 /db_xref="taxon:844"

BASE COUNT 315 a 305 c 267 g 246 t
 ORIGIN

Query Match 99.7%; Score 1129.8; DB 6; Length 1133;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1131: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGGGCGACGAGCCATCATCATCATATGACAGCGCGCTGGTCCGGCGGAGCCAT 60
DB 1 ATGGGCGACGAGCCATCATCATCATATGACAGCGCGCTGGTCCGGCGGAGCCAT 60
QY 61 ATGGCTGACATGATGCTGGGACAGCAAAATGGGTGCGGATCCAGGCGCTCTGTTTGATG 120
DB 61 ATGGCTGACATGATGCTGGGACAGCAAAATGGGTGCGGATCCAGGCGCTCTGTTTGATG 120
QY 121 GCTAAACCCCAAGTACTATCTATCTAGCCACAGAGAGGACACCATCGTGGTGGGGAATCT 180
DB 121 GCTAAACCCCAAGTACTATCTATCTAGCCACAGAGAGGACACCATCGTGGTGGGGAATCT 180
QY 181 AGCGTCAAGAGTACTACTCTGCTGAGACAGTCAACCGTTGATAGCTTTGACAGCGTTC 240
DB 181 AGCGTCAAGAGTACTACTCTGCTGAGACAGTCAACCGTTGATAGCTTTGACAGCGTTC 240
QY 241 CCTGCGATCAACGACCTAGCCACCATCAAGGGGTGAAGAGATCTCAAGATTTGGCTCCCA 300
DB 241 CCTGCGATCAACGACCTAGCCACCATCAAGGGGTGAAGAGATCTCAAGATTTGGCTCCCA 300
QY 301 GAGATGACGGGTAAAGGTGGCTTAACTAGCCAAAGGCTGTCAATGAGCTCTGCGCCAA 360
DB 301 GAGATGACGGGTAAAGGTGGCTTAACTAGCCAAAGGCTGTCAATGAGCTCTGCGCCAA 360
QY 361 AAAGAGACCGGAAGCGGTGATCATCAACCATGAGACATGACACCATGAGAGAGACCGTTTC 420
DB 361 AAAGAGACCGGAAGCGGTGATCATCAACCATGAGACATGACACCATGAGAGAGACCGTTTC 420
QY 421 TTCTCTCAACCTCAGCGGTGAAAAAGCCAAACCTGTCTGCTGTAGAGCGCCATGCTGCA 480
DB 421 TTCTCTCAACCTCAGCGGTGAAAAAGCCAAACCTGTCTGCTGTAGAGCGCCATGCTGCA 480

```

```

Db      421 TTCTCAACCTCAGCGTGAACCAAAACCTGTCCTGTTAGCGCCATGCTCA 480
QY      481 GGCCTTCATGAGTGTGATGGCCCATGAAATCTATAGCGCGTGAATGAGGATC 540
Db      481 GGCCTTCATGAGTGTGATGGCCCATGAAATCTATAGCGCGTGAATGAGGATC 540
QY      541 AACCAAGCTCTACTACAAAGAGTGTGATGATGATGATGATGATGATGATGATG 600
Db      541 AACCAAGCTCTACTACAAAGAGTGTGATGATGATGATGATGATGATGATGATG 600
QY      601 AGAGAAGGACCACTCAACACCAACCGAGTCAATGCTTTGCTTCCGACACAGT 660
Db      601 AGAGAAGGACCACTCAACACCAACCGAGTCAATGCTTTGCTTCCGACACAGT 660
QY      661 AAAATGGCAGAGTCTATATAGCAAGTCAATGCTTCACTCAATGCTTCACTCA 720
Db      661 AAAATGGCAGAGTCTATATAGCAAGTCAATGCTTCACTCAATGCTTCACTCA 720
QY      721 ACCCTGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db      721 ACCCTGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY      781 TAGCTACACCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db      781 TAGCTACACCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
QY      841 GGAATCATCATGCAAGGATGCAAGTGAACCTTCCCTTGAATCAAAATGCTCT 900
Db      841 GGAATCATCATGCAAGGATGCAAGTGAACCTTCCCTTGAATCAAAATGCTCT 900
QY      901 GAAACACAGCAATATGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db      901 GAAACACAGCAATATGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY      961 ACCACCAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Db      961 ACCACCAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1020
QY      1021 AACCTCAAAAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db      1021 AACCTCAAAAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY      1081 GCGATCCAAAGATCTCTCACCCTATTAATCCAAAGAGGATGCTGCTGCT 1133
Db      1081 GCGATCCAAAGATCTCTCACCCTATTAATCCAAAGAGGATGCTGCTGCT 1133

RESULT 3
LOCUS      WSANSAGEN                      1035 bp      DNA      linear      BCT 01-APR-1995
DEFINITION W.succinogenes ansa gene.
ACCESSION  X83689
VERSION     X83689.1 GI:758651
KEYWORDS   ansa gene; asparaginase.
SOURCE      Wolinella succinogenes
ORGANISM    Wolinella succinogenes
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
            Helicobacteraceae; Wolinella.
REFERENCE   1
AUTHORS     Derst, C. and Roehm, K.H.
TITLE       Cloning and sequencing of L-asparaginase from Wolinella
            succinogenes
JOURNAL     Unpublished
AUTHORS     Roehm, K.H.
TITLE       2 (bases 1 to 1035)
JOURNAL     Submitted (05-JAN-1995) K.H. Roehm, Physiologische Chemie der
            Philipps-Universität, Karl-von-Frisch-Str.1, D-35033 Marburg(Lahn), FRG
FEATURES
            source
            1..1035
            /organism="Wolinella succinogenes"
            /mol_type="genomic DNA"
            /db_xref="taxon:844"

```

```

gene      /clone_l1b="lambda EMBL3"
           20..1012
           /gene="ansa"
           20..1012
CDS
           /gene="ansa"
           /EC_number="3.5.1.1"
           /codon_start=1
           /transl_table=11
           /product="L-asparaginase"
           /protein_id="CAA58658.1"
           /db_xref="GI:758652"
           /db_xref="SWISS-PROT:P50286"
           /translation="MAKPYTILATGTTAGSGESSVKSYSAGAVTDKLAAPAI
           NDLATIKGEQISSISQEMTGVWLAKRVNELLAOKTEAVIITHGDTMEETFA
           LNEIVRSOKPVLVAVASMPGSSADGPNLNVAVAIINKASTNGVYIVNDETHA
           AEARLNTAVNAFASPTGKIGTVYKVEYFTQSVPHLASFDISKTEELPRV
           DILVAHPDIDVLVNAALOAGAKGIHAGMGNPEPLTONLEKRAKSGYVASSR
           VSGSGTQAEVDKRIKGFVATESINPQARVILMLALTRKIDREAIKIFSTY"
BASE COUNT  293 a 277 c 236 g 229 t
ORIGIN
Query Match      90.7%; Score 1027.8; DB 1; Length 1035;
Best Local Similarity 99.8%; Pred. No. 1e-284;
Matches 1029; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
103 AGCGCTGCTGTTGATGCTTAACCCCAAGTATCTTACGACAGAGGACCAATC 162
Db      5 AGCGCTGCTGTTGATGCTTAACCCCAAGTATCTTACGACAGAGGACCAATC 162
QY      163 GCTGTTGGGGGAATCTAGCGTCAAGAGTACTCTGCTGCTGAGGACGTCACCTGAT 222
Db      65 GCTGTTGGGGGAATCTAGCGTCAAGAGTACTCTGCTGCTGAGGACGTCACCTGAT 124
QY      223 AAGCTTCTTCAAGCCGTCCTGCAATCAACGACCTTACGACATCAAGGATGACAGATC 282
Db      125 AAGCTTCTTCAAGCCGTCCTGCAATCAACGACCTTACGACATCAAGGATGACAGATC 184
QY      283 TCAAGCATTTGGCTCCCAAGAGATGACGGGTAAAGTGTGGCTTAAACTGACCAAGCGTTC 342
Db      185 TCAAGCATTTGGCTCCCAAGAGATGACGGGTAAAGTGTGGCTTAAACTGACCAAGCGTTC 244
QY      343 AATGAGCTCTGCCCCCAAAAGAGACCGGATGATCATCACCAATGGAAGTACAC 402
Db      245 AATGAGCTCTGCCCCCAAAAGAGACCGGATGATCATCACCAATGGAAGTACAC 304
QY      403 ATGGAAGAGACCGCTTCTTCTCAACTGACGTTGAAGACCAAAACCTGTCCTT 462
Db      305 ATGGAAGAGACCGCTTCTTCTCAACTGACGTTGAAGACCAAAACCTGTCCTT 364
QY      463 GTAGCGGCATGCTCCAGGCTCTTCATGATGATGATGATGATGATGATGATGAT 522
Db      365 GTAGCGGCATGCTCCAGGCTCTTCATGATGATGATGATGATGATGATGATGATG 424
QY      523 GCCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 582
Db      425 GCCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 484
QY      583 GATGAGATTCACGGCGGACAGAGGACCAAGTCAACACCAACCGGATGATGATG 642
Db      485 GATGAGATTCACGGCGGACAGAGGACCAAGTCAACACCAACCGGATGATGATG 544
QY      643 GCTTGGCCCAACACAGGTAAATGCGACAGTCTATATGCAAAAGTGAATTTACT 702
Db      545 GCTTGGCCCAACACAGGTAAATGCGACAGTCTATATGCAAAAGTGAATTTACT 604
QY      703 CAATCGTTCGACTTCACACCTTTCGACAGTGTGATGATGATGATGATGATGAT 762
Db      605 CAATCGTTCGACTTCACACCTTTCGACAGTGTGATGATGATGATGATGATGAT 664
QY      763 CCGAGATCGATATTTCTTACGCTCAACCCGATGATGATGATGATGATGATGATG 822
Db      665 CCGAGATCGATATTTCTTACGCTCAACCCGATGATGATGATGATGATGATGATG 724

```

QY 823 CTTCAGGAGAGAGCCAAAGAAATCATTCATGAGGATGGGCAATGGGAACCTTTCCCT 882
 |||||||
 DB 725 CTTCAGGAGAGAGCCAAAGAAATCATTCATGAGGATGGGCAATGGGAACCTTTCCCT 784
 |||||||
 QY 883 TTGACTCAAAATGCTCTTGAAGAAAGAGCCAAATTCAGGCTAGTCTGCTCGCAAGCTCT 942
 |||||||
 DB 785 TTGACTCAAAATGCTCTTGAAGAAAGAGCCAAATTCAGGCTAGTCTGCTCGCAAGCTCT 844
 |||||||
 QY 943 AGAGTGGGAGTGTGTCCACCAACCAAGAGCTGAAGTGATATAAGAACTTGTTT 1002
 |||||||
 DB 845 AGAGTGGGAGTGTGTCCACCAACCAAGAGCTGAAGTGATATAAGAACTTGTTT 904
 |||||||
 QY 1003 GTGGCTACAGAGTGTCAACCTCAAAAAGCCAGAGTGTCTTATGTAGCCCTCAC 1062
 |||||||
 DB 905 GTGGCTACAGAGTGTCAACCTCAAAAAGCCAGAGTGTCTTATGTAGCCCTCAC 964
 |||||||
 QY 1063 AAAAGTGTGTGTGTGTCCACCAACCAAGAGCTGTCTTCAACCTATTAATCAAGAAAGG 1122
 |||||||
 DB 965 AAAAGTGTGTGTGTGTCCACCAACCAAGAGCTGTCTTCAACCTATTAATCAAGAAAGG 1024
 |||||||
 QY 1123 AATCTCTTCAAC 1133
 |||||||
 DB 1025 AATCTCTTCAAC 1035
 |||||||

RESULT 4
 WSDCUANSA
 LOCUS
 DEFINITION
 W.succinogenes dcua and ansa genes.
 X89215
 X89215.1 GI:895917
 ansa gene; asparaginase; C4-dicarboxylate membrane transporter;
 keywords
 dcua gene.
 succinogenes
 Wolinella succinogenes
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 Helicobacteriaceae; Wolinella.

SOURCE
 ORGANISM
 1
 Lubkowski, J., Palm, G.J., Gilliland, G.L., Derst, C., Rohm, K.H. and
 Wiedower, A.
 TITLE
 Crystall structure and amino acid sequence of Wolinella succinogenes
 L-asparaginase
 JOURNAL
 Eur. J. Biochem. 241 (1), 201-207 (1996).
 MEDLINE
 97054610
 PUBMED
 8898907
 REFERENCE
 2 (bases 1 to 2505)
 Roehm, K.H.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (27-JUN-1995) K.H. Roehm, Institut fuer Physiologische
 Chemie, Philipps-Universitaet Marburg, Karl-von-Frisch-Str.1, 35033
 Marburg, FRG
 COMMENT
 Related sequences X83689.
 FEATURES
 source
 location/Qualifiers
 1..2505
 /organism="Wolinella succinogenes"
 /mol_type="genomic DNA"
 /db_xref="taxon:844"
 /clone_lib="lambda EMBL3"
 1..1068
 /gene="dcua"
 <1..1068
 /gene="dcua"
 /codon_start=1
 /transl_table=11
 /product="C4-dicarboxylate membrane transporter"
 /protein_id="CAA61502.1"
 /db_xref="GI:895918"
 /db_xref="SWISS-PROT:Q34245"
 /translation="KILRSQPKHTEFLAVVYFWTLTLAGTGHAFSLPIAKVAKE
 QGVKSVLETCRGRSLATITSPYSAATVPMGSGRDRLGVTGLLALICIPPTFG
 FYATATISKNVGLIOMPVIGRDSALIMKMLTTATLIMPKIDTDVAVKASTFRSGMTA
 CIGVAVMLADTPEVSNHINDIKALSGILIEHPMMLIILFASMLLSQAATAKAL

misc_feature
 1069..1188
 /note="intergenic region"
 terminator
 1079..1102
 /note="potential rho-independent terminator for dcua"
 1171..1177
 /note="putative"
 1189..2181
 /gene="ansa"
 1189..2181
 /gene="ansa"
 /EC_number="3.5.1.1"
 /EC_number="3.5.1.1"
 /codon_start=1
 /transl_table=11
 /product="asparaginase"
 /protein_id="CAA61503.1"
 /db_xref="GI:895919"
 /db_xref="SWISS-PROT:P50286"
 /translation="NAKPVTLIATGTTAGSGESSVKSYSAGAVTVKLAAPAI
 NDLATIKGEOLISIGSOEMTGRVWLKLAIRVVELLAQKTEAVITIHGDTMEIAF
 LNLATVRSQKPVVLYGAMRPGSSMSADPMLNVAHVNAINKASTNGVVIYVMDDELHA
 AREATKINTAVNAFASPMYKIGVYKVEYFTQSVBPHLASFEDLSKIDELRV
 DILVAHPDPTDVLVNAALQAGAKGIHAGMGNNPPLIONALEKAKSGVVAARRSR
 VSGSTTORAEVDKKLGFVATESLNPKARVIMLALITKTSREAIQKIFSTY"
 2186..2219
 /note="potential rho-independent terminator for ansa"
 BASE COUNT
 594 a 690 c 553 g 668 t
 ORIGIN

Query Match 90.7%; Score 1027.8; DB 1; Length 2505;
 Best Local Similarity 99.8%; Pred. No. 1.1e-284;
 Matches 1029; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 103 AGGCGCTGTGTTTGTGCTAAACCCCAAGTACTATCTTACGACAGAGGACCATC 162
 |||||||
 DB 1174 AGGCGCTGTGTTTGTGCTAAACCCCAAGTACTATCTTACGACAGAGGACCATC 1233
 |||||||
 QY 163 GCTGTTCGGGGGAATCTGCTCAAGAGTACTCTGCTGAGCAGTACCGTTGAT 222
 |||||||
 DB 1234 GCTGTTCGGGGGAATCTGCTCAAGAGTACTCTGCTGAGCAGTACCGTTGAT 1293
 |||||||
 QY 223 AAGCTTCTGACACCGCTCCCTGCATCAACGACCTAGCCACCATCAAGGTTGAACATC 282
 |||||||
 DB 1294 AAGCTTCTGACACCGCTCCCTGCATCAACGACCTAGCCACCATCAAGGTTGAACATC 1353
 |||||||
 QY 283 TCAAGCATTCGCTCCCAAGAGATGAGGGTAAAGTGGCTTAAGTACCAACGCTGTC 342
 |||||||
 DB 1354 TCAAGCATTCGCTCCCAAGAGATGAGGGTAAAGTGGCTTAAGTACCAACGCTGTC 1413
 |||||||
 QY 343 AATGAGCTCTCGCCCAAAAAGAGACCGAAGCCGTGATCATCACCATGAGACTGACAC 402
 |||||||
 DB 1414 AATGAGCTCTCGCCCAAAAAGAGACCGAAGCCGTGATCATCACCATGAGACTGACAC 1473
 |||||||
 QY 403 ATGGAAGACCGCTTCTTCTTCAACCTCAACGCTGAGAAACCAACCAACCGTCCCTT 462
 |||||||
 DB 1474 ATGGAAGACCGCTTCTTCTTCAACCTCAACGCTGAGAAACCAACCAACCGTCCCTT 1533
 |||||||
 QY 463 GTAGGGCCCATGCTCCAGGCTTCCATGAGTGTGAGGCCCATGATCTCTATAC 522
 |||||||
 DB 1534 GTAGGGCCCATGCTCCAGGCTTCCATGAGTGTGAGGCCCATGATCTCTATAC 1593
 |||||||
 QY 523 GCCGTGAATGAGGATCAACAAAGCCTTACTAACAAGAGTGTGATGTGTAAC 582
 |||||||
 DB 1594 GCCGTGAATGAGGATCAACAAAGCCTTACTAACAAGAGTGTGATGTGTAAC 1653
 |||||||
 QY 583 GATAGATTCAGCGCCGACAGAGAGCGCAACGCTCAACACACACCGCATCATTCATT 642
 |||||||
 DB 1654 GATAGATTCAGCGCCGACAGAGAGCGCAACGCTCAACACACACCGCATCATTCATT 1713
 |||||||
 QY 643 GCTTCGCCCAACAGATTAATGCGCAGTCTATTATGGCAAGTGGAGTATTTCAT 702
 |||||||
 DB 1714 GCTTCGCCCAACAGATTAATGCGCAGTCTATTATGGCAAGTGGAGTATTTCAT 1773
 |||||||

QY	703	CAATCGGTGCACCTGCACACCCCTTGCAAGTGAATTGATATAGCAAAATCGAACAATC	762	
Db	1774	CAATCGGTGCACCTGCACACCCCTTGCAAGTGAATTGATATAGCAAAATCGAACAATC	1833	
QY	763	CCAGAGTGCATATTCCTTACGCTCACCCCGATGATACATGATGTTTAACTCAATGCAGCC	822	
Db	1834	CCAGAGTGCATATTCCTTACGCTCACCCCGATGATACATGATGTTTAACTCAATGCAGCC	1893	
QY	823	CTTCAGGCAGAGCCAAAGAAATCATCATCATGCGAGATGGGCAATGGGAACCTTCCCT	882	
Db	1894	CTTCAGGCAGAGCCAAAGAAATCATCATCATGCGAGATGGGCAATGGGAACCTTCCCT	1953	
QY	883	TTGACTCAAAATGCTCTTGAAGAAAGCAGCCAAATCAGGCGCTAGCTGCTGCAAGCTCT	942	
Db	1954	TTGACTCAAAATGCTCTTGAAGAAAGCAGCCAAATCAGGCGCTAGCTGCTGCAAGCTCT	2013	
QY	943	AGAGTGGGCACTGCTCCACCCACCCAGAGGCTGTAAGTGAATATAGAAACCTTGGTTTT	1002	
Db	2014	AGAGTGGGCACTGCTCCACCCACCCAGAGGCTGTAAGTGAATATAGAAACCTTGGTTTT	2073	
QY	1003	GTGCTACAGAGAGCTCCACCTCAAAAGCCAGATGCTCTTATGTTAGCCCTCAC	1062	
Db	2074	GTGCTACAGAGAGCTCCACCTCAAAAGCCAGATGCTCTTATGTTAGCCCTCAC	2133	
QY	1063	AAACTAGTGTATGAGAGGCGCATCCAAAGATCTTTCACCTTTATATCAGAAAGGG	1122	
Db	2134	AAACTAGTGTATGAGAGGCGCATCCAAAGATCTTTCACCTTTATATCAGAAAGGG	2193	
QY	1123	AATCTCTTCAC	1133	
Db	2194	AATCTCTTCAC	2204	
RESULT 5	MSAJ2933	3606 bp	DNA linear	BCT 05-OCT-2000
LOCUS	MSAJ2933	3606 bp	DNA linear	BCT 05-OCT-2000
DEFINITION	Molinitella succinogenes aspa, dcau genes and partial ansa gene.			
ACCESSION	AJ002933			
VERSION	AJ002933.1	GI:2644958		
KEYWORDS	ansa gene; aspa gene; aspartate ammonia-lyase; C4-dicarboxylate membrane transporter; dcau gene; L-asparaginase.			
SOURCE	Molinitella succinogenes			
ORGANISM	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Molinitella.			
REFERENCE	1			
AUTHORS	Ullmann, R., Gross, R., Simon, J., Unden, G. and Kroger, A.			
TITLE	Transport of C(4)-dicarboxylates in Molinitella succinogenes			
JOURNAL	J. Bacteriol. 182 (20), 5757-5764 (2000)			
MEDLINE	20461222			
PUBMED	11004174			
REFERENCE	2 (bases 1 to 3606)			
AUTHORS	Ullmann, R.			
TITLE	Direct Submission			
JOURNAL	Submitted (19-NOV-1997) Ullmann R., Institut fuer Mikrobiologie, Johann-Wolfgang-Goethe Universitaet, Biozentrum N240, Marie-Curie-Strasse 9, 60439 Frankfurt, GERMANY			
COMMENT	Related sequence: X89215.			
FEATURES	Location/Qualifiers			
SOURCE	1..3606			
	/organism="Molinitella succinogenes"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:844"			
	400..1815			
	/gene="aspa"			
	400..1815			
	/gene="aspa"			
	/codon_start=1			
	/transl_table=11			
	/product="aspartate ammonia-lyase"			
	/protein_id="CAA05764.1"			
	/db_xref="GI:2644959"			
gene				
CDS				

[illegible]

DB 3595 GTAGGCCCATG 3606

RESULT 6
LOCUS AP005085/c 303450 bp DNA linear BCT 05-MAR-2003
DEFINITION Vibrio parahaemolyticus DNA, chromosome 2, complete sequence, 2/6.
ACCESSION AP005085 BA000032
VERSION AP005085.1 GI:28808465
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

1
Nasu, H., Iida, T., Sugahara, T., Yamachi, Y., Park, K.S., Yokoyama, K.,
Makino, K., Shinagawa, H. and Honda, T.
A filamentous phage associated with recent pandemic Vibrio
parahaemolyticus O3:K6 strains
J. Clin. Microbiol. 38 (6), 2156-2161 (2000)
20295086
10834969

2
Makino, K., Oshima, K., Kurokawa, K., Yokoyama, K., Uda, T.,
Tagomori, K., Iijima, Y., Najima, M., Nakano, M., Yamashita, A.,
Kubota, Y., Kimura, S., Yasunaga, T., Honda, T., Shinagawa, H.,
Hattori, M. and Iida, T.
Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
distinct from that of V. cholerae
Lancet 361 (9359), 743-749 (2003)
22508454
12620739

3 (bases 1 to 303450)
Oshima, K., Kurokawa, K., Makino, K., Yokoyama, K., Yasunaga, T.,
Honda, T., Shinagawa, H., Hattori, M. and Iida, T.
Direct Submission
Submitted (09-APR-2002) Ken Kurokawa, Osaka University, Genome
Information Research Center; 3-1, Yamadaoka, Suita, Osaka 565-0871,
Japan (E-mail: kengen-info.osaka-u.ac.jp,
URL: http://www.gen-info.osaka-u.ac.jp/, Tel: 81-6-6879-8365,
Fax: 81-6-6879-2047)
genome project
This clone was isolated from a patient presenting with acute
gastroenteritis.

FEATURES
source
location/Qualifiers
1. 303450
/organism="Vibrio parahaemolyticus"
/mol_type="genomic DNA"
/strain="O3:K6"
/sub_strain="RIMD 2210633"
/db_xref="taxon:670"
/chromosome="2"
81..1574
/gene="VPA0300"
81..1574
/gene="VPA0300"
/note="similar to EMB:CA6187.1 (A1591787) percent
identity 68 in 487 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical signal peptide protein"
/protein_id="BAC61643.1"
/db_xref="GI:28808466"
/translation="MNLKAVPKLLVTLASSIPSSVAYAAPNANESSLVQOQEGN
WPTQNEAEKLYOELIYORALSAITQKQPMANVIGMDGSEKFGKGNVLPIMKRD
SRTLVPTPNADVTSNNTLDLKEGLVVAAPKVIQFTDFEORTITDVGILGDRRA
RGGLVLLPNDVGVKPVKGYFTFTSTNNVLFETVMKFGNGADSPAVVIAQOTR
IPLWDEKNIQPMKFPNAGORINMMYPTDSYEMKQFGNDEPISASISELGGVL
ASIVGKPFNPTVEKQLEKAVTAKMTAORTLGRKDEBNILYKRDQYERAMA
GATSEYQESYLDIDORAAFQYVSSAPAMWRTMNGSKITFTVRSKGDIYNGSH
QYKHLHPANPAKLEAVTLYNVVGCTMPETISQLPSKNGFDNIENKNDGSDIDYFGP
QLPNAGPESNYIKITIGRDLTAIRLYGIGIEFDQTKWPPDVLEKIK"

gene
CDS
complement(1835..2314)
/gene="VPA0301"
complement(1835..2314)
/gene="VPA0301"
/note="similar to DBJ:BA843758.1 (AP003138) percent
identity 57 in 158 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative peptide methionine sulfoxide reductase"
/protein_id="BAC61644.1"
/db_xref="GI:28808467"
/translation="MMEKITYAGGCLMGQEPFMRHLPGVISTEAGRANKTDNQSEY
DGYAACVTRTEPDSQVSIETLMGYPFEIIDPYSIKQGDGGERYRGRYSRDRHLE
IAKSFIAERYDDDKIAVEVLPKYNVPSDEEDQRLTFPPNDYCHITPLDLHKYRK"
2493..2882
/gene="VPA0302"
2493..2882
/gene="VPA0302"
/note="similar to GB:AAA64344.1 (U16865) percent identity
52 in 129 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC61645.1"
/db_xref="GI:28808468"
/translation="MKIEHVAIWERLEELKGEYKFNNAVNDKYNPKRHSYFL
DFESGARLELMSMEGTTCENSHMOYTGIAHFAFGSDQAVDQITKLYEDGYORI
DGPRTYDGYEVCVLDPDGRIELTV"
2974..3714
/gene="VPA0303"
2974..3714
/gene="VPA0303"
/note="similar to GB:AAL20622.1 (AE008775) percent
identity 54 in 243 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative regulatory protein"
/protein_id="BAC61646.1"
/db_xref="GI:28808469"
/translation="MNPQNEILQVNDKRRVQVELSDIIGSVGTIRQDINPLEQ
GYLKRVGATATALQSDIDITRLEVRDITQTLANKRADVARNETVILEGSANLLA
RTLAERGDTVITTPSAVIAHLIRNTSANIILIGVYQHGSEISGPIKLTENTHFS
TAVLIGIDGFDHODGFTSRDMRADIAEATILAKRRNIVLTDSKFGQITPSSIGKTN
ISVLTDRKAPKSDLEOKKIGEVVIG"
complement(3733..4089)
/gene="VPA0304"
complement(3733..4089)
/gene="VPA0304"
/note="similar to GB:AAF6321.1 (AE004376) percent
identity 76 in 117 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC61647.1"
/db_xref="GI:28808470"
/translation="MHOEKLNVYEGTPNIGATKAPPEHYFGQFVDYGPDYAEG
QGLDGFYSAEVQSOTSGALLLVFYSSDHATIEKAKRGGQILRIFEPGGCRPH
FTPEGNEFAVWSESH"
4319..5860
/gene="VPA0305"
4319..5860
/gene="VPA0305"
/note="similar to REF:NP_231225.1 (NC_002505) percent
identity 83 in 510 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="catalase"
/protein_id="BAC61648.1"

gene	71142	CCCCAAGTACCTAGCCACAGAGGACACATCCCTGGTTCGGGGGAAATCTAGCGTC	186
CDS	71142	CCTAATATCAAAATCTCGCGACTGGCGACATAGCTGGCGAGGTCAATCTGCCACT	71083
gene	187	AAGAGTAGCTACTGCTGGACAGCTACCGTGTGATAGCTTCTTGACAGCCGCTCCGCC	246
CDS	71082	GAGTCAGACTACACCGCAGGTAAAGTGGCGTGAATATTAATTTCTGCTGCGCGAGC	71023
gene	247	ATCAAGACCTAGCCACCATCAAGAGGTGTAACAGATCTCAAGCATTTGGCTCCCAAGAGATG	306
Query Match	28.5%	Score 323.2; DB 1; Length 303450;	
Best Local Similarity	58.0%	Pred. No. 3.4e-81;	
Matches	571; Conservative	0; Mismatches 413; Indels 0; Gaps 0;	
gene	71142	CCCCAAGTACCTAGCCACAGAGGACACATCCCTGGTTCGGGGGAAATCTAGCGTC	186
CDS	71142	CCTAATATCAAAATCTCGCGACTGGCGACATAGCTGGCGAGGTCAATCTGCCACT	71083
gene	187	AAGAGTAGCTACTGCTGGACAGCTACCGTGTGATAGCTTCTTGACAGCCGCTCCGCC	246
CDS	71082	GAGTCAGACTACACCGCAGGTAAAGTGGCGTGAATATTAATTTCTGCTGCGCGAGC	71023
gene	247	ATCAAGACCTAGCCACCATCAAGAGGTGTAACAGATCTCAAGCATTTGGCTCCCAAGAGATG	306

Db	71022	ATGACCAACATTCGCGATATTTCTGGTGGACAGATCGATCGGCTCTCAACATG	70963
Db	307	ACGGTAAAGTGGCTTAACTAGCCAGCTGTCAATGAGCTCTGCCCCAAAAGG	366
Db	70962	AACGATAGGTATGCTGAAAGTTAGGAAAAGGTTAAGAGTGTCTCAAGTATG	70903
Db	367	ACCAAGCGGTGATCACCACATGAAACATGAAAGAGACCGCTTCTCTC	426
Db	70902	GTTGACGTATGCTATCACCACGATGACGATACGTTGGAATAAAACCGATCTCT	70843
Db	427	AACCTCAGGTGAAAAGCCAAAACCTGCTGCTTGTAGCGCCATGCTCGAGCTCT	486
Db	70842	GATTAAGGTGAAAGAGGACAGACCGCTGTGATCTGCGCGGAGGACCATCACC	70783
Db	487	TCCATGAGTGTGATGCGCCCATGAAATCTATACGCGCGATGTAGCATCAACAA	546
Db	70782	GCCATGAGTGGCGAGGCGCGGCTGATATTAACATGCTGTGTACTGCAACGATGA	70723
Db	547	GCTCTACTAACAAAGAGTGGTGTATGTATGAACGATGATTCACGCGCGAGAA	606
Db	70722	GATCTAAAGGCGCGCGGCTGCTAGTGGCAATGATGACATTTGTGATGCGGAGAT	70663
Db	607	GCGACCAAGCTCAACACACCGCATGATGCTTGTGCGCCACAGAGTAAATC	666
Db	70662	GTAACGAAACCAACACACCGCTGATGATGATGCTTCAATGCTTCCAACTGGCGCTCT	70603
Db	667	GCGACATCTATTTATGCAAGTGGATGATTTCTCAATCCGTTGACCTGACACCTT	726
Db	70602	GCTTACATTCACAAACAGCGATGCAAAAGTACCAAGAGGCTGAAAGTAAACACCA	70543
Db	727	GCAAGTAGTGTGATATTTACCAAAATCGAAGAACTCCCGAGAGTCAATTTACCT	786
Db	70542	GAACCCCATTCGACGATCAAAACACTCAACACTACTAAATAGGATTTGATTAAC	70483
Db	787	CACCCGATGATGATGATGATTTGATGATGATGATGATGATGATGATGATGATG	846
Db	70482	TATGGAATGGGTGAGTGTCCAGTCAAAAGCTTGTGATGATGATGATGATGATG	70423
Db	847	ATTCATGCAAGCAGTGGCAATGAGGACCTTGTGATGATGATGATGATGATGATG	906
Db	70422	GTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	70363
Db	907	GCAGCCAAATCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	966
Db	70362	GCGAGCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	70303
Db	967	CAGAGGCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG	1026
Db	70302	TGAGCGCGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1086
Db	1027	CAAAAAGCCAGAGTGTGCTTATGATGATGATGATGATGATGATGATGATGATG	1086
Db	70242	CAAAAAGCCAGAGTGTGCTTATGATGATGATGATGATGATGATGATGATGATG	1086
Db	1087	CAAAAAGTCTTCTCCACCTTTAA	1110
Db	70182	CAGAAGATGTTCCAGTACTTAA	70159

RESULT 7
 AE015308/c 10893 bp DNA linear BCT 18-OCT-2002
 LOCUS Shigella flexneri 2a ser. 301 section 271 of 412 of the complete
 DEFINITION
 genome: genome: AE005674
 AE015308.1 GI:24053358
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Shigella flexneri 2a str. 301
 Shigella flexneri 2a str. 301
 Bacteria: Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Shigella.
 1 (bases 1 to 10893)

AUTHORS

TITLE

Jin, Q., Yuan, Z. H., Xu, J. G., Wang, Y., Shen, Y., Lu, W. C., Wang, J. H.,
Liu, H., Yang, J., Yang, F., Ou, D., Zhang, X. B., Zhang, J. Y., Yang, G. W.,
Wu, H. T., Dong, J., Sun, L. L., Xue, Y., Zhao, A. L., Gao, Y. S., Zhu, J. P.,
Kan, B., Chen, S. X., Yao, Z. J., He, B. R., Chen, R. S., Ma, D. L.,
Qiang, B. O., Wen, Y. M., Hou, Y. D., and Yu, J.
Genome sequence of *Shigella flexneri* 2a: insights into
pathogenicity through comparison with genomes of *Escherichia coli*
K12 and O157
Nucleic Acids Res. 30 (20), 4432-4441 (2002)

JOURNAL
PUBMED
REFERENCES
AUTHORSTITLE
JOURNAL

Submitted (08-MAY-2001) Microbial Genome Center of Chinese Ministry
of Public Health, 100 Yangxin Jie, Xuanwu Qu, Beijing 100052, P.R.
China

FEATURES
source

Location/Qualifiers
1. 10893
/organism="Shigella flexneri 2a str. 301"
/mol_type="genomic DNA"
/strain="301"
/serotype="2a"
/db_xref="taxon:198214"
188. 541
/locus_tag="SF2947"
188. 541
/locus_tag="SF2947"
/note="Residues 1 to 114 of 117 are 80 pct identical to
residues 1 to 114 of a 216 aa protein from *Klebsiella*
pneumoniae gd: AAF06334.1"
/codon_start=1
/transl_table=11
/product="orf, partial conserved hypothetical protein"
/protein_id="AA044428.1"
/db_xref="GI:24053359."
/translation="MKMSRTSLVCAISLAVSTGVNAERVAIKLHNERSHVH
OSPEELAKVEKQISGNNVIRITPSSOMGNARETWELLONGALDMTGSADSEDFN
IYATVLPFLSCFW"
558. 1151
/locus_tag="SF2948"
558. 1151
/locus_tag="SF2948"
/note="Residues 2 to 197 of 197 are 48 pct identical to
residues 145 to 341 of a 341 aa protein from *Bacillus*
halodurans ref: NP_241567.1"
/codon_start=1
/transl_table=11
/product="orf, partial conserved hypothetical protein"
/protein_id="AA044429.1"
/db_xref="GI:24053360."
/translation="MDSTKDGFPALSAVYAGRTSYAKPKTPKPDGLKLRVOPS
PTTIKMLMGSGSPPTISGEVYTAQOGVDAENNVPSMOTRIETAKVESDEH
AKIDPVLSTKTMKRTPEQDQILETAARKSEAYQOKLMEKIDADTRAQAKAMGGEI
KVVKAPRRKAVPFLDFDKKPKRAALLKEFDNAQ"
1352. 1699
/gene="dctQ"
/locus_tag="SF2949"
1352. 1699
/locus_tag="SF2949"
/gene="dctQ"
/locus_tag="SF2949"
/note="Residues 1 to 108 of 115 are 75 pct identical to
residues 22 to 129 of a 136 aa protein from *Klebsiella*
pneumoniae gd: AAF01413.1"
/codon_start=1
/transl_table=11
/product="putative small integral C4-dicarboxylate
membrane transport protein DctQ"
/protein_id="AA044430.1"
/db_xref="GI:24053361."
/translation="MMIVGILGAAYTVGAORHLSDLFALALNKKRQGLSTIVNL

gene
CDS

gene
CDS

gene
CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

ILFAGSVITGGLKIDKTLATTSQVSAMQIPMGVITLPLSGVMMFYALCFING
SIOLKOPADQAS
complement(1722.. 2117)
/gene="insB"
/locus_tag="SF2950"
complement(1722.. 2117)
/gene="insB"
/locus_tag="SF2950"
/note="Residues 1 to 131 of 131 are 100 pct identical to
residues 37 to 167 of a 167 aa protein from *Shigella*
flexneri ref: NP_085393.1"
/codon_start=1
/transl_table=11
/product="orf, conserved hypothetical protein"
/protein_id="AA044431.1"
/db_xref="GI:24053362"
/translation="MIVCAEMDEHMGVGAQRQRMFLFYADRIRTVAAHFGERTL
ATLERLSLTLAFEVVWMTDGMPLYESRKGKLVYSKRYTORIERHNLNROHLAR
LGRKLSFSKSVELHDKVICHYLNIRKRYQ"
complement(2144.. 2419)
/gene="insA"
/locus_tag="SF2951"
complement(2144.. 2419)
/gene="insA"
/locus_tag="SF2951"
/note="Residues 1 to 91 of 91 are 98 pct identical to
residues 1 to 91 of a 91 aa protein from *Shigella flexneri*
ref: NP_085394.1"
/codon_start=1
/transl_table=11
/product="orf, conserved hypothetical protein"
/protein_id="AA044432.1"
/db_xref="GI:24053363"
/translation="MASISIRCPSCSATEGVRRNCKTRAGORILCSHCRTWOLOFT
YIASOPGHOKIIDAMNGVGRASARIMGVGINVLNHLNNSGRSR"
2578. 3777
/locus_tag="SF2952"
2578. 3777
/locus_tag="SF2952"
/note="Residues 1 to 395 of 399 are 61 pct identical to
residues 38 to 431 of a 435 aa protein from *Salmonella*
typhimurium LT2 ref: NP_462086.1"
/codon_start=1
/transl_table=11
/product="putative integral membrane protein, possible
transporter."
/protein_id="AA044433.1"
/db_xref="GI:24053364"
/translation="MLMPFWDIAITVSQRLANGIDNPALLAIPFTRAGTLNMSGG
DPAFSTVAVNVSCTIGLIPSNVLIVFSLTAGVSASIPNAGTIPGLIMLAWIV
YTFILSRVGLPLSRATEAOACAKFLDALPSLIIVTVGGILFATGESALAVY
LMGISDPLILITLINIVLIVGIPDMTPAVIFLPIFLPADELGMADPVHFGIMV
ANLCISGLIPPEVSALFVGSISGVATIQHLINPLLFYALVALVAMITVYIPQSLFI
POLGLAM"
complement(3878.. 4804)
/gene="yggM"
/locus_tag="SF2953"
complement(3878.. 4804)
/gene="yggM"
/locus_tag="SF2953"
/note="Residues 1 to 308 of 308 are 99 pct identical to
residues 13 to 320 of a 335 aa protein from *Escherichia*
coli K12 ref: NP_417431.1"
/codon_start=1
/transl_table=11
/product="putative alpha helix chain"
/protein_id="AA044434.1"
/db_xref="GI:24053365"
/translation="WMITGNAMADGEPTEIILKDOFKKOYHILKILDAISILKILDAK
GNQATWSAEGVSSDDIYTWVGLADYELLEDQWTEDKRVKFSALMISKGPASWS
VNFISFOAASDRGRVVDIKTNKKYLIIVNSEDFNRFSGLESALNTQKNSIPALKE

VKALDKOMVAQAQADAYWCKDANGKQMTREARKKHQORDEENKONDSFAFAVKYD
 KEYQPAIAACHQSECEYEYVIOQKRFEDINEQROFLOSKLSRFLQDMDWTLEK
 GQYPLIMKVEISEINKVAIILKIDINOANER"
 complement(4957. .6003)
 /gene="ansb"
 /locus_tag="SF2954"
 complement(4957. .6003)
 /gene="ansb"
 /locus_tag="ansb"
 /locus_tag="SF2954"
 /note="Residues 1 to 348 of 348 are 99 pct identical to
 residues 1 to 348 of a 348 aa protein from *Escherichia*
 coli O157:H7 EDL933 ref: NP_289529.1"
 /codon_start=1
 /transl_table=11
 /product="periplasmic L-asparaginase II"
 /protein_id="AAN4435.1"
 /db_xref="GI:24053366"
 /translation="MEPFRKALALVMSGSAALPNTTILATGCTIAGGDSATK
 SNYTAGKGVENLVNAVPLQDLANVGEQVNVIGSDMDNVLTAKKINADCDRT
 DGEVTHGTDMETAYFLDLVYKCDKPVYAVGAMRSTMSADGPNLVNAVYTAAD
 KASANRGVLYMNDVLDGDDVTKTNTDVATKSVNGPLGYIHNKIDYQFPAK
 HTSDTPDVAKMLPEKVIYVYANASDLPALKVADGDIYASAGVNGNLEKSPF
 DTLTAKNGTAVVRSRPVGAITODAEVDATYGVASGTINPKARVILQALALQ
 TKDPOOIOIFENY"
 complement(6179. .6898)
 /gene="y99n"
 /locus_tag="SF2955"
 complement(6179. .6898)
 /gene="y99n"
 /locus_tag="SF2955"
 /note="Residues 1 to 239 of 239 are 99 pct identical to
 residues 1 to 239 of a 239 aa protein from *Escherichia*
 coli O157:H7 ref: NP_311861.1"
 /codon_start=1
 /transl_table=11
 /product="orf, conserved hypothetical protein"
 /protein_id="AAN4436.1"
 /db_xref="GI:24053367"
 /translation="MKRMLLAALSYTAMTANHDVQSVTPRDVIVSPQVQVKE
 NGNVLPDGNVNVNGQYSLNAAQROADYQAELESTLPMWIDEGAKSVEKRIAL
 DKLIIVPDGSSKMRSLRKLDAQLEKQONRIIETRSDDLPHKALIDQVRAEQOQY
 NQAMGILQDSINEMGAKAVKSGSNPLQNVLSLGLQSSIQTEWKKQKDFQGFQK
 DVCSRVTLEDSRKALVGNLX"
 Query Match 26.5%; Score 300; DB 1; Length 10893;
 Best Local Similarity 57.1%; Pred. No. 9.7e-75;
 Matches 568; Conservative 0; Mismatches 420; Indels 6; Gaps 1;

Db 5588 CGTCACGCTCATGATGAGCGGACAGCGGTCCATTCACCTGTATACCGCGGTAGTACCCGAG 5529
 Qy 539 TCACAAAGCCTCTACTAACAAGAGTGTGATTTGTATGAACGATGATTCACGCGC 598
 Db 5528 CTGATTAAGCCTCCCTAATGCTGTGCTGTGTGTATGAACGACACCGTACTGAGC 5469
 Qy 539 CCAGAGAGCGACCAAGCTCAACACCGGACAGTCAATGATCTTCCTGCCCAACAG 658
 Db 5468 GTCCGATGTACCAACAAACACACGACGATGACGACCTTCAAGCTGTATTAAGTAC 5409
 Qy 659 GTAAATGCGACAGTCTATTATGCAAGACGACATTTTCACTCAATCCGTTACCTC 718
 Db 5408 GTCCCTGTGGATPACATTCACAAGGTATGATTTACCAACGTCACCCGACGCTAAGC 5349
 Qy 719 ACACCTTGCAAGTGTATTTATTTAGCAAAATGGAAGAACTCCCGAGGTGATATTC 778
 Db 5348 ACACGAGGATACCGCATTCATGCTCTAAGCTGAATGAGCTGCGGAAGTCGACATCG 5289
 Qy 779 TTTACGCTCACCGGATGATGATGTTTATGATCAATGACGCGCTTACGACGAGGCA 838
 Db 5288 TTTTATTAAGCTTACGCTTACGCTTCCGCTTAAAGCACTGATAGATGCGGCTATG 5229
 Qy 839 AAGGATCATCATGACGACATGCGCAATGGGCAACCTTTCCTTGAATCAAAATGCTC 898
 Db 5228 ATGCATGCTTATGAGCGCTGTGTGGTATGATGTAACCTGTATAAATCCGTTGACACCC 5169
 Qy 899 TTGAAAGACGCAAAATACAGCGCGTACGCTGCTCCGACGCTTACGAGTGGCAGTGT 958
 Db 5168 TGGACACCGCGCGGAAAGACGACGCTGCAATGCGCTTCCCGCTACCGAGCGGTG 5109
 Qy 959 CCACACCCCAAGAGCTGATGATGATGAATGAACCTGTTTGTGTGCTACAGAGATC 1018
 Db 5108 CTACCACTCAGAGATGCTAATGATGATGATGATGATGATGATGATGATGATGATGATG 5049
 Qy 1019 TCACCTCTCAAAAGCCAGAGTCTTCTTATGTTAGCCCTTACCAAAAGTATGATAG 1078
 Db 5048 TGAACCCGCAAAAGAGCGGCTCTGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4989
 Qy 1079 AGGCAATCAAAAGATCTCTCCACCTATATATC 1112
 Db 4988 AGCAGATTCAGACATTTTCAATCAGTACTAATC 4955
 RESULT 8
 AE016988/c
 LOCUS
 DEFINITION
 Shigella flexneri 2a str. 2457T section 11 of 16 of the complete
 genome.
 ACCESSION
 AE016988 AE014073
 VERSION
 AE016988.1 GI:30042420
 KEYWORDS
 SOURCE
 ORGANISM
 Shigella flexneri 2a str. 2457T
 Shigella flexneri 2a str. 2457T
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Shigella.
 REFERENCE
 1 (bases 1 to 292906)
 Wei, J., Goldberg, M. B., Burland, V., Venkatesan, M. M., Deng, W.,
 Fournier, G., Mayhew, G. F., Plunkett, G. III, Rose, D. J., Darling, A.,
 Mau, B., Perna, N. T., Payne, S. M., Runyen-Janecky, L. J., Zhou, S.,
 Schwartz, D. C., and Blattner, F. R.
 Complete Genome Sequence and Comparative Genomics of *Shigella*
 flexneri Serotype 2a Strain 2457T
 Infect. Immun. 71 (5), 2775-2786 (2003)
 2 (bases 1 to 292906)
 Wei, J., Goldberg, M. B., Burland, V., Venkatesan, M. M., Deng, W.,
 Fournier, G., Mayhew, G. F., Plunkett, G. III, Rose, D. J., Darling, A.,
 Mau, B., Perna, N. T., Payne, S. M., Runyen-Janecky, L. J., Zhou, S.,
 Schwartz, D. C., and Blattner, F. R.
 Direct Submission
 Submitted (13-JUN-2002) Genetics Laboratory, University of
 Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA
 JOURNAL
 PUBMED
 TITLE
 AUTHORS
 REFERENCE

```

FEATURES
source
    1..292906
        /location="Qualifiers"
        /organism="Shigella flexneri 2a str. 2457T"
        /mol_type="genomic DNA"
        /strain="2457T"
        /serotype="2a"
        /db_xref="taxon:198215"
        /complement(119..3487)
        /gene="recC"
        /locus_tag="S3030"
        /complement(119..3487)
        /gene="recC"
        /locus_tag="S3030"
        /function="enzyme: Degradation of DNA"
        /note="residues 1 to 1122 of 1122 are 96.34 pct identical
        to residues 1 to 1122 of 1122 from Escherichia coli K-12 :
        B2822"
        /codon_start=1
        /transl_table=11
        /product="DNA helicase, ATP-dependent dsDNA/ssDNA
        exonuclease V subunit, ssDNA endonuclease"
        /protein_id="AAP18145.1"
        /db_xref="GI:30042421"
        /translation="MLRVYHNRDLVLEALMEFVERERLDPEPEMILVQSGMAQ
        WLOMTLSQKFGIANIDPLPASFIMDFVFLPEIPRESAFNKQSMKMLTLPLPOL
        LEREDFTLRHYLTDSDKRLKQLSKADLPFOQYLVYRDMLAQWETGHLVEGLGE
        AQMOAPIMKALVEYHELGOFRHNRANLVQRTETLESATTCPPGLSPRVEIGISA
        LPVYVLAQALQAIKRIEHLFTMPCRYWMDGRIETDPAVLAKLIRORRHSFEDRLPL
        FRSENKQGLPNSGEDVDGNSILASNGKLRDVTIYLLDSESOELDAPVMDPMDNL
        LHNIQSDILEENKAVAGVNIIEFSRSDNKRFLPDSITTFHCHSPQREVEYVHDR
        LEALMEAPLPPLPRDIDIVMAADISYSPFIQAVGSADRYLPALISDRRARSHPV
        LEAFISLISLPDSRFSEVDVADLVPAARFDITEGLRYLRQWVESGIRWIDID
        DNVELELPATGQTRFGLRLMLIGYAMESAQCEWGLPYDESSGLIAELVGLAS
        LMOINWRGILAEGERLEMLPYCRMLNFPIDPATEAMTLIEOQMAIIIEGL
        GAYGVAPVSTILDPAELQRLDORISQRLAGVNICITLMPMSIPKVCYLGLMD
        GVPYPROLPIGLFDMSPKRGDSRRDDDTITFLALISAQKIXIYISRTSDNS
        ERPSVYVQELIDITIGOSHTLPDEALNCDESEKRVKHLTLTRMPFDDQNTYQGE
        RQSFARWMLPAASQAGKASEVQPLPPTLPETVPLELQRFWAHPVRAFQMRQVFN
        FHTEDSEIPDTEPFLIEGLSRVQINQOLNLAIVDODDERLFRFRAGDLPYGFGE
        IFWETQCEMOQLADRYIACRQPSMEIDACNGVOITGMLPQVPGGLKMRSL
        SVQGMQLMEHLHYCASGNGESRFLFRKDEWFRPLAEOALHYLSQLEIREG
        MSPALVLPESGAMKLCYDAONAMDDSTLOKATKFLQAEVGNMVGEDDDI
        WYQRLKQILPEPMTETIVEOSQFLPLFRNQS"
        /complement(3500..3850)
        /gene="ppcC"
        /locus_tag="S3031"
        /complement(3500..3850)
        /gene="ppcC"
        /locus_tag="S3031"
        /function="putative enzyme: Not classified"
        /note="residues 1 to 116 of 116 are 81.03 pct identical to
        residues 1 to 107 of 107 from Escherichia coli K-12 :
        B2823"
        /codon_start=1
        /transl_table=11
        /product="prepilin peptidase dependent protein C"
        /protein_id="AAP18146.1"
        /db_xref="GI:30042422"
        /translation="MSASLKNQOGSLPEVAVVAVLWLVMLVLMVITVLSGIQR
        TSNMSLASRNOYQOLMRHGMQOTOLRAISPANMQVNMQTSQACVSIISVILVSPG
        REGMTRLHCPNCO"
        /complement(3835..4200)
        /gene="ygdB"
        /locus_tag="S3032"
        /complement(3835..4200)
        /gene="ygdB"
        /locus_tag="S3032"
        /note="residues 16 to 121 of 121 are 96.22 pct identical
        to residues 16 to 121 of 121 from Escherichia coli K-12 :
        B2824"
        /codon_start=1
        /transl_table=11
        /product="hypothetical protein"

```

```

        /protein_id="AAP18147.1"
        /db_xref="GI:30042423"
        /translation="MLIVGLSLLIQMSQODRSEFASRVSMESQSLRQALVQSLAMG
        KWHMQTQPAVCCSOTRGTDAQVCLMLDNLNELLTAGIEGVSIMRTGEVIDSKTIFS
        PRGMSDFCLKGCALCOLP"
        /complement(4239..4802)
        /gene="ppdB"
        /locus_tag="S3033"
        /complement(4239..4802)
        /gene="ppdB"
        /locus_tag="S3033"
        /function="putative enzyme: Not classified"
        /note="residues 1 to 187 of 187 are 98.39 pct identical to
        residues 1 to 187 of 187 from Escherichia coli K-12 :
        B2825"
        /codon_start=1
        /transl_table=11
        /product="prepilin peptidase dependent protein B"
        /protein_id="AAP18148.1"
        /db_xref="GI:30042424"
        /translation="MPVKEQGFSLLEVLIMAISSVLLGAEFLPALQRESYSTRK
        LALEDEIMLVRFVAKHQLQAGYCHSGCEGLEIYGOGCIIYQMDANGIWDREP
        VKESDQIGFRKHEVLETLRGATSCSGKGMKXTNPDAIITDFQYVRODVSGSPVL
        TVNMRASSEPGTVVDASTYGFNL"
        /complement(4793..5263)
        /gene="pdaA"
        /locus_tag="S3034"
        /note="point mutation"
        /pseudo
        /complement(5447..6241)
        /gene="thyA"
        /locus_tag="S3035"
        /complement(5447..6241)
        /gene="thyA"
        /locus_tag="S3035"
        /function="enzyme: 2-Deoxyribonucleotide metabolism"
        /note="residues 1 to 264 of 264 are 99.62 pct identical to
        residues 1 to 264 of 264 from Escherichia coli K-12 :
        B2827"
        /codon_start=1
        /transl_table=11
        /product="thymidylate synthetase"
        /protein_id="AAP18149.1"
        /db_xref="GI:30042425"
        /translation="MKQYLELMOKYVLDGKNDKRGTTLSIFGHOMRFLQDGFPL
        VTKRCHLRSTIHELMFLDGDITNATVHNNTTIDPBNADENDQDLPYGVKQRRAP
        TPDGRHIDQITVNLNDKNDPSRRIIVSAMNGVELKMAIPACHAFQFVADGKLS
        COLYORSQDVELLPENIASVYALVHMAOCCDLEVDLPWTGDTLHYSNMDOHTL
        QLSREPRPLKLIIRKRESIFDYRFEDFEIEGYDPHPYKAPVAI"
        /complement(6248..7123)
        /gene="lgt"
        /locus_tag="S3036"
        /complement(6248..7123)
        /gene="lgt"
        /locus_tag="S3036"
        /function="enzyme: Macromolecule synthesis, modification:
        phospholipids"
        /note="residues 1 to 291 of 291 are 95.87 pct identical to
        residues 1 to 291 of 291 from Escherichia coli K-12 :
        B2828"
        /codon_start=1
        /transl_table=11
        /product="phosphatidylglycerol-proliprotein
        diacylglycerol transferase; a major membrane phospholipid"
        /protein_id="AAP18150.1"
        /db_xref="GI:30042426"
        /translation="MTSSYLHPEFEDVIFSGVALHMYGLMYLGVGFIFAMULATR
        ANRPSGWTIKNEENLIVAGFLVGLGRIGYVLYNFPQPMADPLVLRWDGMSF
        HGLIGIVIMVIFPARRKSEFQVSDFLALIPFGAGRLGNFINELGRVDPNF
        PFAMIFRSRTEDILLQTNPOWOSIFDTGTGVLRHPSQLEYELLEGVLFILNLYI
        RKPRMGAVSGPLIGAGARILVIEFRQPAQRTGAMVQYIISGQILSIMIYAGVI
        MMWAVRRSPQOHVS"
        /complement(7274..9520)

```

gene


```

/gene="ansb"
/function="enzyme; Degradation of small molecules: Amino acids"
/note="Residues 1 to 348 of 348 are 98.85 pct identical to residues 1 to 348 of 348 from Escherichia coli K-12 Strain MG1655: B2957"
/codon_start=1
/transl_table=11
/product="periplasmic L-asparaginase II"
/protein_id="AAG58088.1"
/db_xref="GI:12517506"
/translation="MEFFKTRRLALVWGSALALPNTITLTAATGTTAGGDSATK
SNITGKXGVEMLVNAVPOLKDIAVNGEÖVVNISÖDMNDVWLTLAKKITDDCKT
DGEVTHGTVDIETETAYFLDLIVKCDKPVVWVGARPSMSADGPFNLVNAVVAAD
KASANRGVLVYVMDNPTVLDGRDVTKNTITVDATFKSVNGPLGYINNGKIDYQRTARK
HSDTPDPVSKINELPKYGIYVNYANASDLPAKALVDAGDGIYASAGVNGRLYSVE
DLTAAKNGTAVVSSSRVPGATGTQDAEVDAAKXGYTASGTLNQKARVLLQALTLQ
TKDPQIOQIENOV"
/complement(1311..2030)
/gene="y99N"
/note="synonym: 24303"
/complement(1311..2030)
/gene="y99N"
/function="orf; Unknown function"
/note="Residues 1 to 239 of 239 are 99.58 pct identical to residues 1 to 239 of 239 from Escherichia coli K-12 Strain MG1655: B2958"
/codon_start=1
/transl_table=11
/product="orf; hypothetical protein"
/protein_id="AAG58089.1"
/db_xref="GI:12517507"
/translation="MMKMLAALSTVTAHADYOCSTPRDDVYSPOTVOYKE
NGNVTTPDGVMTNGKQYSLNAAÖREKADYQAELRSTLPWIDEGASREKARIAL
DKITIVÖGESKMSRSLTKLDAQLKQÖMNRITFRSDGLTFHYKALIDVYXEGQIV
NÖAMGGLÖDSINEMGAKAVLGSÖGNPLÖNVLSGLGÖSSIQTEWKKÖQÖFOFGK
DVCÖSVYTLÖDSRÖALVGNLK"
/complement(2214..2570)
/gene="y99L"
/note="synonym: 24304"
/complement(2214..2570)
/gene="y99L"
/function="orf; Unknown function"
/note="Residues 1 to 118 of 118 are 99.15 pct identical to residues 1 to 118 of 118 from Escherichia coli K-12 Strain MG1655: B2959"
/codon_start=1
/transl_table=11
/product="orf; hypothetical protein"
/protein_id="AAG58090.1"
/db_xref="GI:12517508"
/translation="MGLNVRGEIEMAKNRSRLKKMHIDFÖFELGFSVAPRPEGS
EÖIIDKVDÖFNEVIEPKLAFDGSYTLAMEGLICQÖEIGKTEEHÖALVRKWLÖER
NLDÖVRSSELFÖDWMÖD"
/complement(2540..3259)
/gene="y99H"
/note="synonym: 24305"
/complement(2540..3259)
/gene="y99H"
/function="orf; Unknown function"
/note="Residues 1 to 239 of 239 are 99.58 pct identical to residues 1 to 239 of 239 from Escherichia coli K-12 Strain MG1655: B2960"
/codon_start=1
/transl_table=11
/product="orf; hypothetical protein"
/protein_id="AAG58091.1"
/db_xref="GI:12517509"
/translation="MRNDVISPÖFENGÖRPLRISÖFVRÖGÖRLTGÖEHÖALÖNÖVY
MGVÖFSEDMÖLÖFALÖREÖVTLÖIGFÖGÖSLVÖAMÖRÖDÖFLÖGTEVHSÖPGVÖA
CLASÖHEGÖSLNÖRVMÖCHDÖVÖLÖHKMÖIDÖNSÖRMÖLÖPPÖDÖMÖKÖRÖHÖKÖRÖIÖVÖP
FÖELVÖSKÖIGÖVÖFMÖATÖMÖPYÖAHÖMÖIEVÖMSÖIDÖYKÖNLÖSÖNDYÖVÖPPÖASÖPVÖTK
FÖDÖRÖGÖRLÖHÖGÖVMDÖLÖMÖFÖRYK"

```

```

gene
3420..4472
/gene="muly"
/note="synonym: 24306"
3420..4472
/gene="muly"
/function="enzyme; Macromolecule synthesis, modification: DNA - replication, repair, restr./modification"
/note="Residues 1 to 350 of 350 are 99.42 pct identical to residues 1 to 350 of 350 from Escherichia coli K-12 Strain MG1655: B2961"
/codon_start=1
/transl_table=11
/product="adenine glycosylase; G.C --> T.A transversions"
/protein_id="AAG58092.1"
/db_xref="GI:12517510"
/translation="MÖASÖFSAÖVLDWYDKRGKTLPMÖIDKTPYKVMÖLSEVLMÖQTÖQ
VÖATVYFÖFERÖMÖRPÖTYÖDLÖANÖLÖDÖVÖLÖHMTÖGLÖGYÖARÖNLÖHÖKÖAÖVÖATÖLÖHG
KÖPÖTEFÖEYÖALÖPGÖVÖSTÖGÖALÖLÖSLÖGÖKÖHÖPÖLDÖGVÖNÖRÖVÖLÖASÖVÖSGÖPKÖKEY
ENÖKÖTÖSÖÖVÖTPÖAVÖGERÖNÖMÖDÖGMÖTÖRÖKÖPÖCSÖLÖCÖPÖNGÖTÖAÖTÖNÖNSÖL
YÖPKÖKÖTÖLÖPÖRTÖGTÖFÖLÖHÖDEÖVÖLÖAÖRPÖSÖLÖMGÖLÖCYÖFPÖFÖADEÖSÖIRÖMÖLÖAÖ
RÖIÖADÖNLÖQÖLÖFÖAÖPHÖTÖSHÖHÖDÖIVÖPMÖLÖPVÖSÖFTÖGÖMÖDÖSÖNÖLÖMÖYÖLÖADÖPPÖSVÖGLA
ÖPÖERÖLÖQÖLÖRTÖGÖAPY"
4500..4775
/gene="y99X"
/note="synonym: 24307"
4500..4775
/gene="y99X"
/function="orf; Unknown function"
/note="Residues 1 to 91 of 91 are 100.00 pct identical to residues 1 to 91 of 91 from Escherichia coli K-12 Strain MG1655: B2962"
/codon_start=1
/transl_table=11
/product="orf; hypothetical protein"
/protein_id="AAG58093.1"
/db_xref="GI:12517511"
/translation="MSÖRTFÖCTÖLÖRÖEÖGÖDÖFÖLYÖPÖELÖKÖRÖIYÖNEÖSÖKÖEÖMÖQÖH
ÖKÖYMLÖINÖEKÖLÖMÖNÖAÖEHÖKÖLÖEÖMÖNÖFÖLÖEÖGÖEÖVÖHÖEGYÖTÖPÖDK"
4837..5919
/gene="m1tc"
/note="synonym: 24308"
4837..5919
/gene="m1tc"
/function="putative enzyme; Cell envelop. Murein sacculus, peptidoglycan"
/note="Residues 1 to 360 of 360 are 99.72 pct identical to residues 1 to 360 of 360 from Escherichia coli K-12 Strain MG1655: B2963"
/codon_start=1
/transl_table=11
/product="membrane-bound lytic murein transglycosylase C"
/protein_id="AAG58094.1"
/db_xref="GI:12517512"
/translation="MMÖKÖYÖLÖALÖIÖAPLÖTÖSÖTÖTKÖGÖYÖNÖEÖMÖVÖDÖNGÖFÖLÖMÖQ
FÖAHÖNIÖENÖGÖFÖEÖVÖIÖAGÖPKÖYÖVÖTÖYÖTÖRÖSÖIÖNÖDÖGTÖITÖETÖIÖAGÖEPÖAHÖNR
ÖLÖITÖKÖLÖMGÖDÖPSÖVDÖLÖXÖSDVÖDTÖITÖKEÖFÖLÖTÖÖVÖDÖNGÖPÖLÖMÖGÖRÖASÖNÖFÖADÖYL
KNÖRLÖKSÖNÖGLÖRÖITÖIÖYÖTÖIÖMÖVÖNÖHÖDÖKÖRÖHÖKÖIÖGÖMÖRÖÖSKÖYÖGÖVÖDESÖLÖILÖMÖTE
SÖSENPÖVÖASÖSDÖALÖGMÖQÖVÖHÖÖAGÖNÖFÖSÖÖGÖSÖTÖSÖFLÖDÖPÖASÖNDÖTÖGÖAYÖLA
MÖNÖNVÖLÖGÖIDÖNÖPTÖSRÖRYÖAVÖITÖYÖNGÖGÖSVÖLÖRVÖSÖNDÖIÖAÖNIÖINTÖMÖPGÖVÖYÖTL
TÖTRÖPÖSÖAESÖRRÖYÖLYÖVÖTÖAQÖSÖYÖRR"
6073..7377
/gene="nupc"
/note="synonym: 24309"
6073..7377
/gene="nupc"
/function="transport; Transport of small molecules: Nucleosides, purines, pyrimidines"
/note="Residues 1 to 434 of 434 are 99.76 pct identical to residues 1 to 434 of 434 from Escherichia coli K-12 Strain MG1655: B2964"
/codon_start=1
/transl_table=11
/product="transport of nucleosides, permease protein"

```

/protein_id="AAG58095.1"
 /db_xref="GI:12517513"
 /translation="KCFKHSASRCEETNMNKLKLTILSFLOELCMLXMTLLGYS
 MVTTLKEDGASIGAVSSUGIAVEMPTLINTISYRLONAGMDIYDDPPIRINCTGFI
 AOYTPPEAMLVILINSFAMPTLGLINTISYRLONAGMDIYDDPPIRINCTGFI
 MAMVVSISGEFESMOLYIGALISAIVLFTLPLIPAKQAQMSMTLLGDAF
 ALFKNRKAFEFIESMLGAELOITNNFGLFSLDKDPMFASSTVOHASTIS
 OISEPLITLIPPLSLRYSIKYNNMISIVMILRFAVAGDPTGVLVLSMTIV
 GCAPDFNIGSVFEEKESVPARIBASOGCEIILGIGIVSGKVEYMTONG
 ITDMQVWLIFAGISVYLARAFYAMERIKYRVPITGQYTS"
 complement(7427..9622)

Query Match 26.2%; Score 296.8; DB 1; Length 10749;
 Best local similarity 56.9%; Pred. No. 8.1e-74;
 Matches 566; Conservative 0; Mismatches 422; Indels 6; Gaps 1;

119 TGGCTAAACCCCAAGTACTTCTTACCCACAGAGGACCATCGTTCGGGGCAAT 178
 1074 TGGCATACCCCAATATCCATTTTACACCGCGGGACCATTCGGGTGGTGGTACT 1015
 119 CTAGGCTCAAGTAGTACTCTGCTGAGCAGTACCGCTTGATTAAGCTTCTGCAAGC 238
 1014 CCGCAACCAATCTTACTACACAGGGGTAAGTTGGGTAAATCTGGTTAATGCGG 955
 239 TCCCTGCATCAAGCACTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCC 298
 954 TCCCGCACTGAAGACATTTGGCAAGTTAAAGGAGACAGTAGTGAATTCGGCTCC 895
 299 AAGCATGACGGGTGAAGGTGGCTTAACTAGCCACGAGTCAATGAGTCTCTGCCC 358
 894 AAGCATGACGATGATGCTGCTGACACTGGCCAAAAATTA-----CACCGACT 841
 359 AAAAAAGACCGAAGCGGTGATCATCCCATGGAATGACACCATGGAAGAGACGGCT 418
 840 GCGATTAACCGACGGCTTCTGCTATCTACGAGTACCGACGATGAGGAAGAACGCTT 781
 419 TCTTCTCAACCTCAAGGTGAAGAAAGCAAAACCTGCTGCTGAGCGGCGCATGCTC 478
 780 ACTTCTCGACCTGACGGGTGAAGTGGACAAACCGGTGGTGAATGCTGCGCAATGCTC 721
 479 CAGGCTTTCATGAGTGTGATGAGCCCATGATCTCTATACGCGCTGATGATGAGCA 538
 720 CGTCCAGCTATGAGGAGCAGAGCTCAATCAACCTGATTAACGGGATGATGACCGCAG 661
 539 TCAACAAGCCTTACTACTAACAAGAGGTGATGATGATGATGATGATGATGATGATG 598
 660 CTGATTAAGCTTCCGTAATCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
 599 CAGAGAGAGCAGCAGCTCAACACACGCGCATGATGATGATGATGATGATGATGATG 658
 600 GTGCGCATGTCACCAAAACACACACGCGCATGATGATGATGATGATGATGATGATG 541
 659 GTAAATTCGACAGTCTATTAATGCAAGAGTGAATGATGATGATGATGATGATGATG 718
 540 GTCCGTGGGATACATTCACACGATGATGATGATGATGATGATGATGATGATGATG 481
 719 ACACCTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 778
 480 ACACCAAGGACAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 421
 779 TTTACGTCACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 838
 420 TTTATTAAGCTTAACGATCGGCTTCCGGCTAAAGCCCTGATGATGATGATGATG 361
 839 AAGGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 898
 360 ATGGCATGTGAGCGCTGTGGTGAAGCGCAACCTGATTAATCCCTGTTTGAACAGC 301
 899 TTTAAAGAGCAGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 958
 300 TGGCGACCGCGCGGAGAAAGAGTACTGATGATGATGATGATGATGATGATGATGATG 241
 959 CCACCAACCAAGAGGCTGAAGTGAATGAAGAACTTGTGTTGTGCTACAGAGATG 1018

RESULT 10
 AP002563/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

1019 TCACCTCTCAAAACCCAGAGTGTCTTATGTTAGCCCTACCAACCAATGATAGAG 1078
 180 TCACCTCTCAAAACCCAGAGTGTCTTATGTTAGCCCTACCAACCAATGATAGAG 121
 1079 AGCGATCCCAAAAGATCTCTCCACCTATTAATC 1112
 120 AGCAGATCCAGCAGATCTTCAATGATGATGATGATGATGATGATGATGATGATG 87

1019 TCACCTCTCAAAACCCAGAGTGTCTTATGTTAGCCCTACCAACCAATGATAGAG 1078
 180 TCACCTCTCAAAACCCAGAGTGTCTTATGTTAGCCCTACCAACCAATGATAGAG 121
 1079 AGCGATCCCAAAAGATCTCTCCACCTATTAATC 1112
 120 AGCAGATCCAGCAGATCTTCAATGATGATGATGATGATGATGATGATGATGATG 87

AP002563 266658 bp DNA linear BCT 07-MAR-2001
 Escherichia coli O157:H7 DNA, complete genome, section 14/20.
 AP002563.1 GI:13363121
 Escherichia coli O157:H7
 Escherichia coli O157:H7
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Escherichia.
 1 (sites)
 Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C. H., Kimura, S.,
 Kurokawa, K., Ishii, K., Hattori, M., Tatsuno, I., Abe, H., Iida, T.,
 Yamamoto, K., Ohnishi, M., Hayashi, T., Yasunaga, T., Honda, T.,
 Sasaki, C. and Shinagawa, H.
 Complete nucleotide sequence of the prophage VT-2-Sakai carrying the
 verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7
 derived from the Sakai outbreak
 Genes Genet. Syst. 74 (5), 227-239 (1999)
 2 (sites)
 Ohnishi, M., Mura, T., Nakayama, K., Kuhara, S., Hattori, M.,
 Kurokawa, K., Yasunaga, T., Yokoyama, K., Makino, K., Shinagawa, H. and
 Hayashi, T.
 Comparative analysis of the whole set of rRNA operons between an
 enterohemorrhagic Escherichia coli O157:H7 strain and an
 Escherichia coli K-12 strain MG1655
 Syst. Appl. Microbiol. 23 (3), 315-324 (2000)
 3 (sites)
 Yokoyama, K., Makino, K., Kubota, Y., Matanabe, M., Kimura, S.,
 Yutsudo, C. H., Kurokawa, K., Ishii, K., Hattori, M., Abe, H., Iida, T.,
 Yamamoto, K., Hayashi, T., Yasunaga, T., Honda, T., Sasaki, C. and
 Shinagawa, H.
 Complete nucleotide sequence of the prophage VT-1-Sakai carrying the
 Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli
 O157:H7 strain derived from the Sakai outbreak
 Gene 258 (1-2), 127-139 (2000)
 4 (sites)
 Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K.,
 Yokoyama, K., Han, C.-G., Ohtsubo, E., Nakayama, K., Mura, T.,
 Tanaka, M., Tobe, T., Iida, T., Takami, H., Honda, T., Sasaki, C.,
 Ogasawara, N., Yasunaga, T., Kuhara, S., Shiba, T., Hattori, M. and
 Shinagawa, H.
 Complete genome sequence of enterohemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12
 DNA Res. 8 (1), 11-22 (2001)
 5 (bases 1 to 266658)
 Hayashi, T., Kurokawa, K., Makino, K., Yasunaga, T., Shinagawa, H. and
 Ohnishi, M.
 Direct Submission
 Submitted (26-JUN-2000) Ken Kurokawa, Osaka University, Genome
 Information Research Center, 3-1, Yamadaoka, Suita, Osaka 565-0871,
 Japan (E-mail: ken@gen-info.osaka-u.ac.jp)

URL: <http://www.gen-info.osaka-u.ac.jp/>, Tel: 81-6-6879-8365,
 Fax: 81-6-6879-2047)
 genome project.

FEATURES

Location/Qualifiers

1..26658

/organism="Escherichia coli O157:H7"

/mol_type="genomic DNA"

/strain="O157:H7"

/sub_strain="RIMD 0509952"

/db_xref="taxon:83334"

/gene="Ecs3650"

/complement(244..693)

/gene="Ecs3650"

/complement(244..693)

/note="similar to B2790_ECOLI g11789154 percent identity 100 in 149 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37073.1"

/db_xref="GI:13363122"

/translation="MAEIGIFGTMVGNLSVAEAEALITAGGKATPEDEPESDM
 LPYQDRVIVVSTTQGGDLPSIVLFGIKDLSGEOPLRGVATLADSSVAVNCN
 GGRQFDALLQEQSAQRYGEMILIDASENPEPEPESNPWVEQNGTLLS"

/complement(711..1493)

/gene="Ecs3651"

/complement(711..1493)

/note="similar to YQCB_ECOLI g11789155 percent identity 99 in 260 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37074.1"

/db_xref="GI:13363123"

/translation="MEILYODEMLVAVKSGMVLHRSWLDREKVVYVOTVRDQIS
 QHVFTHRLDRPTSGVLLMGLSSEAGRLAQOFEQHOIQKRYHATVRGMLPEAVLDY
 LVEELDKTADFAREDEKPGPAVHYHGLAOFVEMVATGRYPTRVGVLEPRTGR
 KHQRRHLAHLRPIIGDSKGLDRONSGAHEGRLQIMHAASLSLTHPTGEPPL
 IHAGLDITWMAISQFGRGLPENNEREFSAPSQDGERSS"

/complement(1493..1822)

/gene="Ecs3652"

/complement(1493..1822)

/note="similar to B2792_ECOLI g11789156 percent identity 98 in 109 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37075.1"

/db_xref="GI:13363124"

/translation="MTTHDRVRLQDLALREHOMRDEPQHPQFNSTQPFMDT
 MEPEMLQWVLIPEMHDLLNNQDLPFAFVAVPYEALATDHPORALLIALEKIDA
 LFADDA"

/complement(1865..2233)

/gene="Ecs3653"

/complement(1865..2233)

/note="inhibitor of CsrB"

/complement(2444..2989)

/gene="Ecs3653"

/complement(2444..2989)

/note="similar to SYD_ECOLI g11789157 percent identity 98 in 181 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="Syd protein"

gene

CDS

/protein_id="BAB37076.1"

/db_xref="GI:13363125"

/translation="MDLLRQALDKFTARYCDAMHEHNSMPLSEELYGVSPSPCIIST
 TPDAYWQPPPTGEOGNVAVAFVIVQPTIHFTYTQFGMDHAGQDGIKLLLO
 TNSDEDFRRQVEMLIGHLVYQKRLKLPPLFLATLEEFLEVISVNLSEVEKEELGT
 KRRTHLAPLAEEFLNOLKPL"

/complement(3057..3905)

/gene="Ecs3654"

/complement(3057..3905)

/note="similar to YQCB_ECOLI g11789158 percent identity 98 in 282 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37077.1"

/db_xref="GI:13363126"

/translation="WSTYANHOALAGLTGRSTDYRDYDASLLQGVRLNRPDLGL
 KADNLPFGQDITWILYELSMNLAKGLPQAVAGHLELDYTVNLIESKSPKLYXSNFO
 TRFNMDVRYOTLERDLSTCAQGEVSVALYRLDEEGPQIGHNGCTIDODITDYN
 EFTYDLENNATSGEKVVEETLYSHLKNCLITHQPPMGSIQIOYRGRQIDREKLYN
 LVSEFHNEFHQCVETLFDNLRFQCEPKLSYARTTRRGGLDINFRNSDSEVPT
 TRLYVQ"

/complement(4017..5381)

/gene="Ecs3655"

/complement(4017..5381)

/note="similar to YQCB_ECOLI g11789159 percent identity 99 in 454 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="hypothetical protein"

/protein_id="BAB37078.1"

/db_xref="GI:13363127"

/translation="MTHTSPICSMMDLSOLEYMDLKRASSDLYOLFNCISLAVLNS
 GSLDNKSEKLESEFENDIVLRREGVKTLEINPEPAFVDRIRIALLQANLFAVL
 DILEVYQILNIVYRFPPLNDNSVHTNTYELILNRARLHNGEARNMVVCGHGSIN
 ENELVARRVGNGLGRLNLCGCCGGAEMAPKGAAGAQAQRYKDSRFTGMEPS
 IIAEPNRYNELLIMPDIKRLAEVRIAHGIIIFPGVGTAEELVLLGLNMPA
 NKQOVLPLITGKESADYFRVLDDEFVHTLGENARHRIIIDAEEVARRMKKSM
 LVKENRDGTGDAYSFMSMRKIAFDLQMPESPSENNANLKVDPQVEVLADLRRA
 SGIVAGNVEVGIKRAIEFERGYKINDKEIMRMDLLGFAQGRMKRPGSATYPCY
 EICT"

/complement(5938..7227)

/gene="Ecs3656"

/complement(5938..7227)

/note="similar to SDAC_ECOLI g11789160 percent identity 100 in 429 aa (Conserved in E.coli K-12)"

/codon_start=1

/evidence=not_experimental

/transl_table=11

/product="probable serine transporter"

/protein_id="BAB37079.1"

/db_xref="GI:13363128"

/translation="METTOTSTIASKDSRSAMRKTDYMWMLGLYATGACVLPPLIN
 AGVGKIPILIMAILAFPMTEFAHRLGLRPFVLSGNGEDTTEVEEHEFGGAKRLTI
 ILVFAIYPIILVSVALTITVESPMHSHQDLPAPRAISLILVGMNTVREGEOM
 IVKASILVFPFVGLMILALYLIPONGALLETLSIDTASATNGMLMTLMLIPM
 VSEFNHSPILSEFAKREYGDABQKSKILFAHIMVLTVMFVFCVSLTPA
 DLAARKONISLISLANHFNAPVYAMAPILAIATKSLFGLHVGARREGFNGMVK
 SLRGGKSIETKLNRLTALPLMTVTVVATVINSIIGMETLGGPIIAMILFLMPRY
 AIQKVPARKYSGHISNVFVVMGLIASIFYSLS"

/complement(7285..8652)

/gene="Ecs3657"

/complement(7285..8652)

gene

CDS

Query Match

Best Local Similarity 56.9%

Score 296.8; DB 1; Length 26658;

Matches 566; Conservative 0; Mismatches 422; Indels 6; Gaps 1;

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

rene

Query Match	26.2%;	Score 296.8;	DB 1;	Length 303121;
Best Local Similarity	56.9%;	Pred. No. 1.4e-73;		
Matches 566;	Conservative 0;	Mismatches 422.	Indels 5	

[illegible]

```

QY 359 AAAAGAGCCGAGCCGTGATCATCAACCGTGAACCTGACACCATGAGAGACCGCTT 418
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68968 GCGATAAACGTGAGCGCTTCTCTCTTTACCCAGCTACGACACATGGAAGAAACCGCTT 68909
QY 419 TCTTCTCAACCTCAGCGGTGAAAAAGCCAAAACCTGTCTGCTTGTAGCGCCATGCGTC 478
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68968 ACTTCTGACCTGAGCGGTGAAATGCGCAACACCGGTGATGATGCTGATGCAATGCGCC 68849
QY 479 CAGGCTTCCATGAGTCTGATGAGCCCAATGATCTCTTAACAGCCCGTGAATATGCGA 538
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68848 GGTCCACCTCTATGAGCGCAACACCGATTCATTCATGATGATGATGATGATGATGATG 68789
QY 539 TCAACAAGCCCTCTCTCAACAAGAGAGTGTATGATGATGATGATGATGATGATGATGATG 598
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68788 CTGATAAAGCCCTCCGCTTAATGCTGCGCTACGTGATGATGATGATGATGATGATGATG 68729
QY 599 CCAGAGAGAGCAGCAAGCTCAACACCGACGATGATGATGATGATGATGATGATGATGATG 658
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68728 GCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 68669
QY 659 GTAAATGCGCAGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 718
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68668 GTCTCTGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 68609
QY 719 ACACCTTGTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 778
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68608 ACACCGAGAGCAGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 68549
QY 779 TTATGAGCTACCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 838
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68548 TTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 68489
QY 839 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 898
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68488 ATGCGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 68429
QY 899 TGTAAAAAGCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 958
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68428 TTGCAACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 68369
QY 959 CCACCGAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1018
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68368 CTACCGAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 68309
QY 1019 TCAACCTGCAAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1078
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68308 TGAACCGAGAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 68249
QY 1079 AGGCGATGCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68248 AGCAGATGCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 68215

```

```

TITLE Smith, H. O. and Venter, J. C.
JOURNAL Whole-genome random sequencing and assembly of Haemophilus
MEDLINE Influenzae Rd
PUBMED Science 269 (5223), 496-512 (1995)
AUTHORS 95350630
REFERENCE 7542800
TATSOV, R. L., Mushegian, A. R., Bork, P., Brown, N. P., Hayes, W. S.,
Borodovsky, M., Rudd, K. E. and Koonin, E. V.
Metabolism and evolution of Haemophilus influenzae deduced from a
whole-genome comparison with Escherichia coli
CURR. BIOL. 6 (3), 279-291 (1996)
JOURNAL MEDLINE
PUBMED 96398784
REFERENCE 3 (bases 1 to 10173)
AUTHORS White, O., Clayton, R. A., Kerlavage, A. R. and Fleischmann, R. D.
JOURNAL Submitted (25-JUL-1995) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
4 (bases 1 to 10173)
REFERENCE White, O., Clayton, R. A., Kerlavage, A. R. and Fleischmann, R. D.
AUTHORS Submitted (27-SEP-1997) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr, Rockville, MD 20850, USA
REMARK The H. influenzae sequence has been updated by R. Fleischmann. New
database matches have been assigned, product names have been
improved, and a number of frame shifts have been corrected. We
incorporated their annotation into the /notes fields of the
corresponding H. influenzae genes
5 (bases 1 to 10173)
REFERENCE White, O., Clayton, R. A., Kerlavage, A. R., Fleischmann, R. D.,
AUTHORS Peterson, J., Hickey, E., Dodson, R. and Gwin, M.
JOURNAL Direct Submission
REMARK Submitted (28-MAY-1998) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr, Rockville, MD 20850, USA
REMARK The whole genome was shifted by 588 nucleotides for a new start
location/qualifiers
FEATURES
source
gene
CDS
gene
CDS
gene
CDS

```

gene
CDS
/protein_id="AAC22407.1"
/db_xref="GI:1573749"
/translation="MGNLVAECCRCITELNNGWTEVTSVPPTLANNAPYGSVEN
GKISSAPAELEAAADYAAVALVOYGEFGALIAQMDMTSGALPCCIKRWYLY
IKNHLIS"
complement(2249, .2758)
/gene="H10743"
complement(2249, .2758)
/note="similar to SP:P15040 GB:M24489 PID:147797
PID:466747 GB:U00096 percent identity: 56.16; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="protein-export protein (secb)"
/protein_id="AAC22401.1"
/db_xref="GI:1573750"
/translation="MSEOKDVAATEEQPVLOIQRIVYKDVSEAPNLPHTFOQEWK
PKIGFDISTEYVDVLEVINISVETLEDSDGVAFCIEVROAGVFTTSGEDVQ
MAHLISQCPNMLEPYARELYSNLVNNGTEPALMLSPVNEDELFVEYMNROAENAE
KSEEOYRK"
complement(2772, .3218)
/gene="H10744"
complement(2772, .3218)
/note="similar to SP:P37688 PID:466749 GB:U00096
PID:179040 percent identity: 50.74; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC22402.1"
/db_xref="GI:1573751"
/translation="MOEFLSMATEFAOKHTLAVSWFAIFVWVITYFYKATSKFYI
THNEVIRLINSDEAIVDLRSLEPQGHINSINVLPSKIKNONICKLSHKNALI
LVDTNGTSASASAVFLTKQFNSVFLKEGLSAMVAANLPLVKKH"
3444, .4493
/gene="H10745"
3444, .4493
/note="similar to GB:M34234 SP:P08005 GB:M34277 PID:145277
PID:146597 percent identity: 70.52; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="L-asparaginase II (ansb)"
/protein_id="AAC22403.1"
/db_xref="GI:1573752"
/translation="MKLTALCTLEFGVSIANMADLPNITLITATGTIAGSGSSV
NSAYRAGQSIDTLLEAVPEMKNIANIKGEQIVKISGOMNDEWMLKAIINAOCKS
TGFVTHGTDMETAYFLDLTKCEKFPVVGAMRPATERSADGELNLNAYVYRA
DKSSGRGLVAMNNEVIGARDYKSTTAQVDFHSPNGSGCYTINSYVDERSPE
KRTIMPEVVERLDSLPRYGIYAYSNAYPEPLNALLNAGYGTISAGVNGVNAH
LDRLEKKAQDSYVYVSSRVPTGYITTRDAEVDKSYFVASGLTNQKARVLIQALT
QTRKRVYIDQITEDF"
4649, .5971
/gene="H10746"
4649, .5971
/note="similar to GB:U14003 SP:P14409 GB:M27058 GB:X79886
PID:146047 percent identity: 42.59; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="anaerobic C4-dicarboxylate membrane transporter
protein (deub)"
/protein_id="AAC22404.1"
/db_xref="GI:1573753"
/translation="WSAMFLQFAVLICIFGARVGGIGLVEFGGLIALISFEGFL
KPAGLPIDVMEIMAVVAAAAMQAGGIDYMIKATILRNPHITFIIVAAVWLF
TLAGTGVAVSVLIVIAVSRONGIRPERPSMAVIVASQFAIVASPIIAAVAVVAF
LEPOGILGADVLMVITPSTILGLFLACLFVNKMKELDDPEYORLNDPRKADIFAS
TTSYKEVSVSTAKISVSLFLGALLVYLGMASIRIVFDKRFMGMAHTTETIMLSI

gene
CDS
GALLILCKPDGTATTKGSVFRAGNRAVIAIFGLIAMLGDTLMQAHITKVEKGLVE
TAPWAFALFVLVSLVNSOGATVATLEPDLALGIPAPVIGVVAANGVFPIPNIG
PLIASIDEDTGTETIKGYILNHSFMLPGLISMFFCLLIGLSHVL"
6175, .7509
/gene="H10747"
6175, .7509
/note="similar to SP:P00393 PID:581140 GB:U00096
PID:1651546 PID:1787352 percent identity: 57.77;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH dehydrogenase (ndh)"
/protein_id="AAC22405.1"
/db_xref="GI:1573754"
/translation="MKNVYIVGGAGGIEATLFLGNKLRGROKAKVTLVDNATHWK
PLHEIATGVMDVDGVSLSYRAHGNHFESEGSITIRINECKYVLLAPYQGECK
LVIRRIPIYDLVIAIGKSNDEMTKCAVNCIFLDSKALRFOHKLLELFESN
RALDDIGEEPRKQIDENKYNIVAVGGATGVLIELYHAEDLSYGGKIDSSC
LOVTLVAGRTFLPALBENLSAAVLDELEKENGIVQNLMTTEAQPNTLITKDEIRK
ADLVPAAGVASTVTOFQDLEINRINQVLDKDTQVYDSDSIFATGDAALIOSNG
KLVPRAQAHOAKAKACAKNIFLFEKPKLSEFYNDKGLVLSNFTALGSLNFKG
KNPLVQGEKFAQFVYSLYRHHQHALGCIKIGIILVDKLNRFLKRLKH"
complement(7589, .10021)
/gene="H10748"
complement(7589, .10021)
/note="similar to SP:P00482 GB:K00127 PID:147295
PID:148102 PID:148109 percent identity: 57.21; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="glycerol-3-phosphate acyltransferase (plsB)"
/protein_id="AAC22406.1"
/db_xref="GI:1573755"
/translation="MANFIMYRQLSLPLSALVKNPPIANPIEELSLNIHQPIYV

Query Match 26.28; Score 296.4; DB 1; Length 10173;
Best Local Similarity 57.18; Pred. No. 1,1e-73;
Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;
DB 126 ACCCCAGTACTATCTTACCCAGAGGACCCAGTGGTGGGGAATCTAGCGT 185
3515 ACCAAATATTACATCTTGGCAACGGGTGTACCATTCAGAGGCGCAAGTTCGT 3574
186 CAAGAGTAGCTACTCTGTCGAGCAGTCACCGTTGAATACCTCTGACCCCTCTGC 245
3575 AATTTCTGCTATTAAGCTGTGACAAATTAATGATTTGACTTAAATGAAGCTGTACCA 3634
246 CATCAACGACCTTACCCAGGAGTGAACAGATCTCAAGCATGGCTCCCAAGAT 305
3635 AATGAAATATTGCGCAACATTTAAAGGTGAGCAAAATTTGAAAATAGGTTCAACAACAN 3694
306 GACGGTAGGRTGTGCTTAACTAGCCAAAGCTGTCAGTGAAGTCTGCGCCCAAAAAGA 365
3695 GAATGACGAAGTGTGCTTAAAGTGCACAAAGCCATCAATGCTC-----AATGTAAAG 3748
366 GACCGAAGCGTGATCATCCCATGAACTGACACATGAGAGAGACCGCTTCTTCT 425
3749 TACTGATGATTTGTCTATACCATGTGTACAGTATACGAGAGAGACCGCTTATTTCT 3808
426 CAACCCAGCGTGAAGCAAAAACCTGTGCTTGAAGCGCATCGCTCCAGGCTC 485
3809 AGATTTTAACCGTAAATGGAAGAACGGGTGTCTCGTTGGGGCAATCGCTGCAAC 3868
486 TTCCATGATGCTGATGAGCCCATGATCTATTAAGCCGCGTGAATGAGCATCAACA 545
3869 AGAAAAAAGTGTGATGAGCCCATTAATCTTACAAATGCTGTGCTGCGACGACACA 3928
546 ACCCTCTACTAACAAGAGTGTGATGTGATGAACGATTCACCGCCGACAGAGA 605
3929 AAAATCAAGTGTGCTGTGTTTGTGTCGATCAATATATGAAGTACTAGTGTGCGGA 3988

```

QY 606 AGCGACCAAGCTCAACACCCAGCAATGATCTTGGCCCAACACAGTAAT 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3989 TGTACAAAACACAGGACGAGTCCAAACGTTCCATTCACCAATTTATGTTCT 4048
QY 666 CGGACAGCTCTTTATNGCAAGTGAATTTCACTCAATCCCTTCAGACCT 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4049 AGGCTATATTCATTAACAGCAAGTGAATTTCACTCAATCCCTTCAGACCT 4108
QY 726 TCGAAGTGAATTTATTAAGCAAAATGCAAACTCCCAAGTCAATTTCTTACGC 785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4109 CAACACTCGTTTAAAGTGAATAATTAAGTACCTACCAAGTGGGATTTATTAAGC 4168
QY 786 TCACCCCGATGATCTGATGTTTATGATCAATGACCCCTTCAGAGGCAAGGAAT 845
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4169 TTAATCAATGCAACCTGTCGAACCAATTAACGATTAACATGCTGATCAAGGAT 4228
QY 846 CATCATGAGGATGAGCAATGGAACCTTCCCTTTCAGTCAAAATGCTTGAANA 905
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4229 TGTATCTGAGAGTGGCAATGGAATGTAATGCTGACACTTATGATGCTTGAANA 4288
QY 906 AGCAGCCAAATCAAGCGTATGCTGCTGCAAGCTTAAGTGGGAGTTCACAC 965
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4289 AGCCGCAAAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4348
QY 966 CCAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1025
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4349 TGTGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4408
QY 1026 TCAAAAAGCCAGAGTCTTATGATGATGATGATGATGATGATGATGATGATGAT 1085
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4409 ACAGAAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4468
QY 1086 CCAAAAGATCTTCTCCACCTATTA 1110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4469 TCAACATATTTTCAAGACTTTCA 4493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13
AR274513_07

WPCOMMENT
Sequence split into 19 fragments LOCUS AR274513 Accession AR274513

Fragment Name	Begin	End
AR274513_00	1	110000
AR274513_01	100001	210000
AR274513_02	200001	310000
AR274513_03	300001	410000
AR274513_04	400001	510000
AR274513_05	500001	610000
AR274513_06	600001	710000
AR274513_07	700001	810000
AR274513_08	800001	910000
AR274513_09	900001	1010000
AR274513_10	1000001	1110000
AR274513_11	1100001	1210000
AR274513_12	1200001	1310000
AR274513_13	1300001	1410000
AR274513_14	1400001	1510000
AR274513_15	1500001	1610000
AR274513_16	1600001	1710000
AR274513_17	1700001	1810000
AR274513_18	1800001	1910000

Continuation (8 of 19) of AR274513 from base 700001 (AR274513 Sequence 1 from patent US

Query Match 26.2% Score 296.4; DB 6; Length 110000;
Best Local Similarity 57.1%; Pred. No. 1.5e-73;
Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

```

QY 126 ACCCAAGTACTATCTAGCAAGAGAGCAATCGCTGTTGGGGGAATCTAGGT 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102722 ACCCAATATTAATCTAGCAAGAGAGCAATCGCTGTTGGGGGAATCTAGGT 102781
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 186 CAAGACTACTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 102782 AATTTGCTATTAAGCTGAGCAATTAATTAATTAATTAATTAATTAATTAAT 102841
QY 246 CATCAAGAGCTTATGAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102842 AATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 102901
QY 306 GACGGGTAAAGTGTGAGTAACTTAATTAATTAATTAATTAATTAATTAATTAAT 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102902 GAATGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 102955
QY 366 GACCGAAGCCGTATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102956 TACTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 103015
QY 426 CAACCTCAGGTGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 485
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103016 AGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 103075
QY 486 TTCCATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAG 545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103076 AGAAAAAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAG 605
QY 546 AGCCTCTACTAACAAGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAG 605
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103136 AATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 103195
QY 606 AGCGACCAAGCTCAACACCCAGCAATGATCTTGGCCCAACACAGTAAT 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103196 TGTAAACAAAGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAG 103255
QY 666 CGGACAGCTCTTATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAG 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103256 AGCTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 103315
QY 726 TCGAAGTGAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 785
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103316 CAACACTCGTTTAAAGTGAATAATTAATTAATTAATTAATTAATTAAT 103375
QY 786 TCACCCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 845
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103376 TTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 103435
QY 846 CATCATGAGGATGAGCAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 905
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103436 TGTATCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 103495
QY 906 AGCAGCCAAATCAAGCGTATGCTGCTGCAAGCTTAAGTGGGAGTGTTCACAC 965
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103496 AGCCGCAAAAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 103555
QY 966 CCAAGAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1025
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103556 TCGTGAAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 103615
QY 1026 TCAAAAAGCCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1085
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103616 ACAGAAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 103675
QY 1086 CCAAAAGATCTTCTCCACCTATTA 1110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103676 TCAACATATTTTCAAGACTTTCA 103700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 14
AR274513_08
WPCOMMENT
Sequence split into 19 fragments LOCUS AR274513 Accession AR274513

Fragment Name	Begin	End
AR274513_00	1	110000
AR274513_01	100001	210000
AR274513_02	200001	310000
AR274513_03	300001	410000
AR274513_04	400001	510000

AR274513_05 500001 610000
 AR274513_06 600001 710000
 AR274513_07 700001 810000
 AR274513_08 800001 910000
 AR274513_09 900001 1010000
 AR274513_10 1000001 1110000
 AR274513_11 1100001 1210000
 AR274513_12 1200001 1310000
 AR274513_13 1300001 1410000
 AR274513_14 1400001 1510000
 AR274513_15 1500001 1610000
 AR274513_16 1600001 1710000
 AR274513_17 1700001 1810000
 AR274513_18 1800001 1910000
 AR274513_19 1900001 2010000

Continuation (9 of 19) of AR274513 from base 800001 (AR274513 Sequence 1 from patent US

Query Match 26.2%; Score 296.4; DB 6; Length 110000;
 Best Local Similarity 57.1%; Pred. No. 1.5e-73;
 Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

QY 126 ACCCCAGTACTCTCTAGCCACAGAGGACCCATCCCTGTTGCGGGAGATCTAGCAT 185
 DB 2722 ACCAAATATTCACATCTGGACAGAGGCTGTACCATTCAGAGAGCGGCAAGTTCGCT 2781
 QY 186 CAAGAGTAGTACTCTGCTGGAGAGTACCGTGTATAGCTTCTTGACGCGCTCCCTGC 245
 DB 2782 AAATTCGCGTATTAAGCTGACATTAATTAATTAATTAATTAATTAATTAATTAATTA 2841
 QY 246 CATACAGACCTTCCACCATCAAGGAGTGAACAGATCTCAAGATTCGCTCCCAAGAT 305
 DB 2842 AATGAAAAATATTCGCAACATTAAGGTGACCAATTAATTAATTAATTAATTAATTAAT 2901
 QY 306 GAGCGTAAAGTGTGGTAACTAAGCAAGCGTGTCAATAGCTCCTCGCCCAAAAAGA 365
 DB 2902 GAATGAGAGTGTGGTAACTAAGCAAGCGTGTCAATAGCTCCTCGCCCAAAAAGA 2955
 QY 366 GACCGAAGCGGTATCTCAACCATCAAGGAGTGAACAGATTCGCTCCCAAGAT 425
 DB 2956 TACTGATGAGTGTGGTAACTAAGCAAGCGTGTCAATAGCTCCTCGCCCAAAAAGA 3015
 QY 426 CAACCTGAGGTGAAAGCAAGCAAGCGTGTCTGCTGAGCGCCATGCGCTCCAGCTC 485
 DB 3016 AGATTTAAGCGTGAAGTGAAGCAAGCGTGTCTGCTGAGCGCCATGCGCTCCAGCTC 3075
 QY 486 TTCCATGAGTGTGAGCGCCATGAGTGTCTTAAACCGCGTATGATGATCAACAA 545
 DB 3076 AGAAAAAGTGTGATGAGCGCCATGAGTGTCTTAAACCGCGTATGATGATCAACAA 3135
 QY 546 AGCTTACTAACAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 605
 DB 3136 AAATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3195
 QY 606 AGGACCAAGCTTCAACAGCAGCAGTCAATGATGATGATGATGATGATGATGATGAT 665
 DB 3196 TGAACAAAAGCAGTCAACAGCAGCAGTCAATGATGATGATGATGATGATGATGAT 3255
 QY 666 CGGCAAGCTTCAACAGCAGCAGTCAATGATGATGATGATGATGATGATGATGAT 725
 DB 3256 AGGCTATATTCATCAACAGCAGCAGTCAATGATGATGATGATGATGATGATGAT 3315
 QY 726 TGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 785
 DB 3316 CAACACTCGCTTAAAGTGAAGCAATTAAGTGAAGCAATTAAGTGAAGCAATTAAGT 3375
 QY 786 TCACCCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 845
 DB 3376 TTATTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3435
 QY 846 CATCATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 905
 DB 3436 TGTATCTCGAGAGTGTGAGAGAGAGTGTATGATGATGATGATGATGATGATGAT 3495
 QY 906 AGCAGCAATATCAGCAGTGTGCTGCTGAGAGCTTGAAGTGTGAGAGCTTGAAGCAG 965

DB 3496 AACCAGCAAGATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3555
 QY 966 CCAAGAGCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1025
 DB 3556 TGTGACGCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3615
 QY 1026 TCAAAAAGCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1085
 DB 3616 ACAGAAAGCCGCGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3675
 QY 1086 CCAAAAGATCTCTCCACTATTA 1110
 DB 3676 TCACCAATATTTGCAAGACTCTAA 3700

RESULT 15
 ECOLASNII 1530 bp DNA linear BCT 26-APR-1993
 LOCUS E.coli L-asparaginase II (ansb) gene, complete cds.
 DEFINITION M34277.1 GI:146596
 VERSION M34277.1 GI:146596
 KEYWORDS L-asparaginase II.
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Escherichia.
 1 (bases 1 to 1530)
 JENNINGS, M.P. and Beacham, I.R.
 Analysis of the Escherichia coli gene encoding L-asparaginase II,
 ansb, and its regulation by cyclic AMP receptor and FNR proteins
 J. Bacteriol. 172 (3), 1491-1498 (1990)
 PUBMED 90170867
 COMMENT 2407723
 FEATURES
 source
 Original source text: E.coli (strain K-12) DNA.
 location/Qualifiers
 1..1530
 /organism="Escherichia coli"
 /mol_type="genomic DNA"
 /db_xref="taxon:562"
 323..>1431
 /product="ansb mRNA"
 /evidence="experimental"
 385..1431
 /EC_number="3.5.1.1"
 /note="L-asparaginase II (ansb) precursor"
 /codon_start=1
 /transl_table=11
 /protein_id="AA24062.1"
 /db_xref="GI:146597"
 /translation="MEFKKTAALALVMEFGSALALPNTTILATGTTAGGDSATK
 SNTVKGVEENLVNAVPOKDIANVGEVNIIGSDMNDVLTLLAKRINDCKRT
 DGFVTHGDTMETEAVFDLTVKCDKPVVAVGAMPSPMSGDGFENLYNAVTPAD
 KASANGVLYVMNDVLDGSDVTKNTTDAVPSVNGPLGYTHNGKIDYQETPAK
 HESDPEDVSKINELPKVGIIVYANASDLPAKAIYDAGCDGIVSGVNGNLIKSVF
 DLTATAKTGIVAVRSRVPFGATITDPAEDDAKTFVASGITLNPQKARVLQALTO
 TNDPQIQIIFNY"
 385..450
 /note="L-asparaginase II signal peptide"
 451..1428
 /product="L-asparaginase II"
 /EC_number="3.5.1.1"
 /base_count 409 a 381 c 382 g 358 t

sig_peptide
 mat_peptide
 BASE COUNT 409 a 381 c 382 g 358 t
 ORIGIN

Query Match 26.1%; Score 295.2; DB 1; Length 1530;
 Best Local Similarity 56.8%; Pred. No. 1.7e-73;
 Matches 565; Conservative 0; Mismatches 423; Indels 6; Gaps 1;

QY 119 TGGCTAAACCCCAAGTATCTTACCAAGAGGACCAATCGCTGTTGGGGGAGAT 178
 DB 446 TGGATTAACCAATATCAACATTTAGCAACCGGCGGACCAATCGCTGTTGGGACT 505

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2003, 15:11:40 ; Search time 346 Seconds
(without alignments)
8839.489 Million cell updates/sec

Title: US-09-937-982-3

Sequence: 1 atggcagcagcaccatca.....aagaaggaatctcttcac 1133

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_19jun03:*

- 1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
- 25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	1129.8	99.7	1133	20	AA03474
2	1129.8	99.7	1133	21	AA062512
3	300.4	26.5	1044	21	AA082834
4	296.4	26.2	1830121	17	AA042063
5	294	25.9	1848	18	AA06346
6	259.2	22.9	1196	19	AA014328
7	223.2	19.7	2837	8	AA070557
8	215.8	19.0	1014	15	AA068438

9	199	17.6	1174	25	AB280801
10	108.8	9.6	345	24	ABX66196
11	103.6	9.1	543	24	AA521124
12	102	9.0	543	24	AA521116
13	102	9.0	543	24	AA521118
14	102	9.0	543	24	AA521122
15	102	9.0	543	24	AA521126
16	101.6	9.0	1436	25	ACC44571
17	101.4	8.9	552	19	AA013947
18	100.4	8.9	681	24	AA521115
19	100.4	8.9	681	24	AA521121
20	100.4	8.9	681	24	AA521125
21	100.4	8.9	682	24	AA521117
22	100.4	8.9	682	24	AA521123
23	100.4	8.9	776	24	AA521119
24	98.4	8.7	1435	25	ACC44570
25	96.4	8.5	543	24	AA521120
26	94.2	8.3	900	24	AB073014
27	93.4	8.2	1247	20	AA029725
28	82.8	7.3	960	25	ABX07792
29	82.8	7.3	2162598	25	AB56454
30	79.8	7.0	316	25	AB280804
31	75.8	6.7	2365589	24	ABA90521
32	74	6.5	2231	19	AA02398
33	72.6	6.3	366	22	AA025008
34	71.4	6.3	969	22	AA053259
35	71.4	6.3	981	24	AB090934
36	71.4	6.3	4020	22	AA054070
37	71	6.3	360	25	AB280805
38	70.4	6.2	1230	22	AA025127
39	68.4	6.0	6682	20	AA012993
40	68.4	6.0	6682	24	AB598788
41	67.6	5.9	619	21	AA250042
42	67.4	5.9	654	21	AA250043
43	66.6	5.9	501	22	AA066207
44	66.2	5.8	1455	22	AA500250
45	66.2	5.8	2847	22	AA023014

ALIGNMENTS

RESULT 1	
AA03474	
ID	AA03474 standard; DNA; 1133 BP.
XX	
AC	AA03474;
XX	
DT	30-APR-1999 (first entry)
XX	
DE	Mollinella succinogenes L-asparaginase DNA.
XX	
KW	L-asparaginase; amplification; treatment; disease; asparagine depletion;
KW	malignant disease; haematology; lymphoma; leukaemia; myeloma; AIDS;
KW	autoimmune disease; rheumatoid arthritis; systemic lupus erythematosus;
KW	covalent modification; acylation; acylation; pharmacokinetic; immunogenic; spleen;
KW	hypersensitive; reduced toxicity; immunosuppression; allergy; thymus;
KW	lymphocyte marker distribution; hepatotoxicity; ss.
XX	
OS	Mollinella succinogenes.
XX	
PN	W0956410-A1.
XX	
PD	17-DEC-1998.
XX	
PF	09-JUN-1998; 98MO-US11905.
XX	
PR	09-JUN-1997; 97US-0049085.
XX	
PA	(CHIL-) CHILDRENS HOSPITAL LOS ANGELES.
XX	
PI	Durden DL;
XX	

Erwinia carotovora
Helicobacter pylori
CDNA encoding Tm 3
CDNA encoding Tm 2
CDNA encoding Tm 2
CDNA encoding Tm 3
CDNA encoding Tm 7
PN0V4836 glucose 1
Salmonella sefa ge
CDNA encoding Tm 2
CDNA encoding Tm 3
CDNA encoding Tm 7
CDNA encoding Tm 2
CDNA encoding Tm 3
CDNA encoding Tm 1
PN0V4835 glucose 1
CDNA encoding Tm 1
Recombinant SOM95
Insert from pET28C
S. pneumoniae type
Streptococcus pneu
Erwinia carotovora
Genomic sequence o
Streptococcus pneu
Nucleotide sequenc
S. epidermidis ope
Staphylococcus epi
S. epidermidis gen
Erwinia chrysanth
Nucleotide sequenc
Enterococcus faeca
Enterococcus faeca
DNA encoding hepat
DNA encoding hepat
Staphylococcus aur
Lfn-Bcl-XI apoptos
Nucleotide sequenc


```

Db      62 TGGCATTTACCAATTCACATTTTACGACCGGGGACCAATGCGGTGGTGGTACT 121
QY      179 CTAGCGTCAAGAGTACTACTCTGCTGGAGGAGTACAGTGTATAGCTTCTGCAGCGC 238
Db      122 CCGGACCAAAATCTTACTACACAGCGGGTAAAGTTGGCGTGAATAATCTGGTTAAWGGG 181
QY      239 TCCCTGCGATCAAGACCTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCC 298
Db      182 TGGCGCACTAAAGACATTTGGAGACGTTAAAGCGAGACAGGTATGATGATCGGCTCC 241
QY      299 AAGAGATACGGGTAAAGTGTGGCTTAACCTAAGCCAGCGTCAATAGCTCTCGGCC 358
Db      242 AGGACATGAACGATATGTCTGGTACACCTGGCGAAAAAATTA-----CACCGACT 295
QY      359 AAAAGAGACCGAAGCCGTGATCATCACCCATGGAACCTGACACCATGGAAGAGACGGCT 418
Db      296 GCGATTAACCGAGCGGCTTCGTCATTAACCGGATACGACACGATGGAAGAAACCGCTT 355
QY      419 TCTTCTCACTCAAGCTGAGTAAAGCCAAACCTGTCTGCTGTAGCGCCATGCGTC 478
Db      356 ACTTCTCGACCTCAAGCTGAGTAAAGCCAAACCGGTGATGGTGGCGCAATGGCC 415
QY      479 CAGCTCTTCCATAGTGTGATGGCCCATGAACTCTATTAAGCCCTGATGAGCGA 538
Db      416 CGTCCAGCTCCATAGCGGACAGCGTCCATTAACCTGATTAAGCGGTAGTGAACCGAG 475
QY      539 TCAACAAAGCCTTACTAACAAGAGAGTGTGATTTGATGATAGATGATGATGATGATG 598
Db      476 CTGATTAAGCTTCGCTAATGCTGCGGCTGTGATGATGATGATGATGATGATGATG 535
QY      599 CCAGAGAGGAGACCAAGCTGACACACCGGATGATGATGATGATGATGATGATGATG 658
Db      536 GTCCGATGATCAACCAACCAACCGGATGATGATGATGATGATGATGATGATGATG 595
QY      659 GTAAATTCGGACATGATATATGCAAAAGTGTGATTTTACTCAATCCGTTGAGCTC 718
Db      596 GTCTCTGGGATATATCAACAGGTAAAGTGTGATGATGATGATGATGATGATGATG 655
QY      719 ACACCTTGAAGTGTGATTTAGCAAAATGCAAGAACTGCCAGAGTGTGATGATG 778
Db      656 ACACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 715
QY      779 TTTACGCTCAACCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 838
Db      716 TTTATATCTACGCTAAGCATCCGATCTTCGCGTAAAGCATGATGATGATGATG 775
QY      839 AAGGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 898
Db      776 ATGGCATCGTTAGCGGTGTGATGATGATGATGATGATGATGATGATGATG 835
QY      899 TTTAAAAAGAGCAATTCAGGCTAGTCTGCTGATGATGATGATGATGATGATG 958
Db      836 TGGCAACCGCGGAAAAAGCACTGATGATGATGATGATGATGATGATGATGATG 895
QY      959 CCAACACCCAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1018
Db      896 CTACACCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 955
QY      1019 TCAACCTCAAAAGCCAGATGATGATGATGATGATGATGATGATGATGATG 1078
Db      956 TGAACCGCGGAAAAAGCGGCTGCTGATGATGATGATGATGATGATGATGATG 1015
QY      1079 AGGCGATCCAAAGATGATGATGATGATGATGATGATGATGATGATGATG 1098
Db      1016 AGCAGATCCAGCATCTTC 1035

```

RESULT 4
 AAT42063
 ID AAT42063 standard; DNA: 1830121 BP.
 XX
 AC AAT42063;

```

XX      14-SEP-1999 (first entry)
DT
XX      Haemophilus influenzae complete genome sequence.
DE
XX      Genome; bacterium; Haemophilus influenzae; computer readable medium;
KW      expression modulating fragment; regulation; gene expression; vector;
KW      organism; open reading frame; ORF; ds.
OS      Haemophilus influenzae.
XX      WO9633276-A1.
PN
XX      24-OCT-1996.
PD
XX      22-APR-1996; 96WO-US05320.
PF
XX      07-JUN-1995; 95US-0487429.
PR      21-APR-1995; 95US-0426787.
PR      07-JUN-1995; 95US-0476102.
XX
XX      (HUMA-) HUMAN GENOME SCI INC.
PA      (UTRO ) UNIV JOHNS HOPKINS.
XX
XX      Adams MD, Fleischmann RD, Smith HO, Venter JC, White O;
XX      WPI, 1996-485782/48.
XX
XX      Haemophilus influenzae Rd genome recorded on computer readable
XX      medium - useful for identifying commercially important nucleic acid
XX      fragments by homology searching
XX
XX      Claim 1; Page 77.2-77.1091; 1291pp; English.
XX
XX      This sequence represents the complete genome sequence of the bacterium
XX      Haemophilus influenzae strain Rd. The invention relates to a computer
XX      readable medium (CRM) having recorded upon it the complete H. influenzae
XX      nucleotide sequence (1), a representative fragment of (1) or a nucleotide
XX      sequence at least 99% identical to (1). By providing the full-length
XX      genomic sequence in a computer readable form, it is possible to identify
XX      commercially important nucleic acid fragments and expression modulating
XX      fragments (EMFs) of the Haemophilus genome. The EMFs can be used to
XX      regulate the expression of a nucleic acid molecule. Vectors and altered
XX      organisms comprising the predicted ORFs can be used to produce any of the
XX      polypeptide fragments of the H. influenzae Rd genome.
XX
SQ      Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;
Query Match      26.28; Score 296.4; DB 17; Length 1830121;
Best Local Similarity 57.18; Pred. No. 8.1e-81;
Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;
QY      126 ACCCAAGTACTTCTAGCAGAGGACGACCATGCGGTGCGGGGAATGATGACGT 185
Db      802722 ACCAATATATTAATCTTGGCAACGGGTGGTACCATTCAGAGAAACGCAATTTGGT 802781
QY      186 CAGAGTACTCTCTGAGAGCATGACCGTTGATTAAGCTTTCAGACCGCTGCTC 245
Db      802782 AAATTCGCGTATTAACCTGACAAATTAATGATGATTAATTAAGGTGACAGA 802841
QY      246 CATCAAGACCTGCGCAACATCAAGAGTGAAGATGATCAAGATTTGGCTCCCAAGAT 305
Db      802842 AATGAAATAATTTGCCAACATTAAGGTGAGCAAAATTTGAAATTAAGTTTCAAGCAT 802901
QY      306 GAGCGGTAGGTGCTTAATAGCCAGCGTGTCAATGAGCTCTGCCCAAAAGA 365
Db      802902 GAATGAGAGTGTGGCTAAATCGGCAAAAGCATTAATGCTC-----AATGTAAG 802955
QY      366 GACCGAAGCGGTATCATCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 425
Db      802956 TACTGATGATTTGTCTATTACCATGATGATGATGATGATGATGATGATGATGATG 803015
QY      426 CAACCTCAGGTTAAAGCAAAACCAAAACCTGCTGCTGATGAGGCGCATGCGTCCAGGCTC 485

```

```

Db      803016 AATTTAAACCCGTAATAATGTGAATAAACCCGTTGTTCTGTTGGGGCAATGCGTCTCAAC 803075
QY      486 TTCCATGAGTGTGATGCCCCCATGATCTCTATTAACGCCGCGATGTGACCATCAACA 545
Db      803076 AGAAAAAGTGTGATGGCCCATTTAACTTACAAATGCTGCTGCTGCGACACAGACA 803135
QY      546 AGCCTCTACTAACAAGAGAGTGTGATGTGATGAACGATGAGATTCACGCCGCCAGAGA 605
Db      803136 AAAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 803195
QY      606 AGCAGCAAGCTCAACACACACGCGCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTG 665
Db      803196 TGTAAACAAAACAGACAGACCGCATGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 803255
QY      666 CGGACAGCTCTATTTAGCAAGTCAATTTCACTCAATCCGCTGACCTCAACCCCT 725
Db      803256 AGCCTATATTCATTAACGCAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 803315
QY      726 TCGAATGAGTTGATATTAAGCAAAATCGAAGAACCTCCGACAGTGCATATTCCTTAACG 785
Db      803316 CAACACTCGCTTTAAGCTAGAAAAAATTAGTATACCTACCAAAAGTGGGATTTATTATG 803375
QY      786 TCACCCCGATGATCTGATGTTTATGTCATGCAAGCCCTTCAGGACGAGCAAGCAAGAT 845
Db      803376 TATTCAAATGCACTGCTGCAACCATTAAGCATTAACATGCTGCTATCAAGGGAT 803435
QY      846 CATCTGAGAGATGAGCAATGGGAACCTTCCTTCCTGACATCAAAATGCTCTTGAAA 905
Db      803436 TGTATTCAGAGGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 803495
QY      906 AGCAGCAATCAAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 965
Db      803496 AGCCGCAAAAGATAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 803555
QY      966 CCAAGAGGCTGAAGTGAATGAAGAACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1025
Db      803556 TCGTGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803615
QY      1026 TCAAAAAGCAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1085
Db      803616 ACAGAAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 803675
QY      1086 CCAAAAGATCTCTCCACCTATTAA 1110
Db      803676 TCAACATATTTTGAAGACTTTAA 803700

```

RESULT 5
AAT96346
ID AAT96346 standard; cdna; 1848 BP.
XX

```

AC      AAT96346;
XX
XX      08-APR-1998 (first entry)
XX
XX      Chimeric gene containing anti-asparaginase MAb light and heavy chain.
XX
XX      Immunoglobulin; Ig; heavy chain; variable region; murine; human;
XX      asparaginase II; monoclonal antibody; MAb; light chain;
XX      recombinant chimeric polypeptide; ss.
XX
XX      Synthetic.
XX      OS Chimeric - Homo sapiens.
XX      OS Chimeric - Mus sp.
XX
XX      US5686579-A.
XX
XX      11-NOV-1997.
XX
XX      23-MAY-1995; 95US-0447422.
XX
XX      22-JUN-1993; 93US-0081410.
PR

```

```

PR      21-JUN-1988; 88US-0205748.
PR      31-AUG-1992; 92US-0938505.
PR      23-MAY-1995; 95US-0447422.
XX
XX      (HYBR-) HYBRISSENS LTD.
XX
XX      Ramjessingh M, Rothstein A, Shami EY;
XX      WPI; 1997-558200/51.
XX
XX      Self-protecting chimeric polypeptide comprising biologically active
XX      PT sequence and single-chain antibody sequence - has resistance to e.g.
XX      disrupting temperature, presence of proteolytic enzymes, etc.
XX
XX      Example 2; Columns 27-30; 29pp; English.
XX
XX      The present sequence is a chimeric gene containing the cDNA for
XX      CC the immunoglobulin (Ig) heavy and light chain variable regions of a
XX      CC murine anti-asparaginase II monoclonal antibody (MAb), and human
XX      CC asparaginase II. The gene was used in the preparation of a novel
XX      CC recombinant chimeric polypeptide, comprising a 1st region
XX      CC comprising a biologically active domain and another domain
XX      CC containing an epitope, linked via a polypeptide to a 2nd region
XX      CC including a single chain antibody (SCA) having the light and heavy
XX      CC chains of an antibody variable region which specifically binds the
XX      CC epitope in the 1st region. The chimeric polypeptide assumes a
XX      CC conformation in which the SCA is bound to the epitope of the 1st
XX      CC region and protects its biological activity from deactivation by
XX      CC denaturing temperatures or pH conditions, proteolytic enzymes,
XX      CC oxidizing agents or alcohol. The regions of the chimeric
XX      CC polypeptide interact to form a structure analogous to an
XX      CC antibody-antigen complex. A L-asparaginase-SCA fusion protein of
XX      CC the above type has better trypsin resistance than free
XX      L-asparaginase.
XX
XX      Sequence 1848 BP; 435 A; 498 C; 517 G; 398 T; 0 other;
XX
XX      Query Match 25.9%; Score 294; DB 18; Length 1848;
XX      Best Local Similarity 56.3%; Pred. No. 9.7E-82;
XX      Matches 574; Conservative 0; Mismatches 440; Indels 6; Gaps 1;

```


Db 1681 ATTCAAGACTATTCCATCTATTGAT 1708

RESULT 8
AA068438
ID AA068438 standard; DNA; 1014 BP.
XX
AC AA068438;
XX
DT 25-MAR-2003 (updated)
DT 12-JAN-1995 (first entry)
XX
DE Pseudomonas glutaminase gene.
XX
KW Glutaminase; antiviral; virucide; anticancer; cancer therapy;
KW HIV virus; gene therapy; Escherichia coli; ds.
XX
OS Pseudomonas sp.
XX
PN MO9413817-A1.
PD 23-JUN-1994.
PF 04-DEC-1992; 92WO-US10421.
PR 04-DEC-1992; 92MO-US10421.
PR 04-DEC-1992; 92AU-0032358.
XX
PA (MEME-) ME MEDICAL ENZYMES AG.
XX
PI Freeman AG, Macalister TW, Roberts J, Sethuraman N;
XX
DR WPI: 1994-217891/26.
DR P-PSDB; AAR59739.
XX
PT Recombinant glutaminase derived from Pseudomonas 7A - expressed
PT in E. coli to increase yield and avoid Pseudomonas endotoxins for
PT antiviral and anticancer therapy
XX
PS Disclosure: Page 33; 60pp; English.
XX
CC Chromosomal DNA from Pseudomonas sp. 7A (ATCC 29538) was used to
CC construct a genomic library in Escherichia coli LB392. Screening
CC with mixed oligonucleotide probes was used to isolate a glutaminase-
CC encoding clone. This was sequenced using the primers given in
CC AA068439-47. The gene can be used to manufacture recombinant
CC glutaminase, free of Pseudomonas exotoxin, for use in e.g. HIV and
CC cancer therapy. The gene may also be used in gene therapy protocols.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 1014 BP; 226 A; 324 C; 318 G; 146 T; 0 other;

Query Match 19.0%; Score 215.8; DB 15; Length 1014;
Best Local Similarity 52.6%; Pred. No. 3.3e-57;
Matches 519; Conservative 0; Mismatches 462; Indels 6; Gaps 2;

QY 128 CCCAAGTACTATCTCCAGCAGAGGACCATCGCTGGGGGAATCTAGCTCA 187
DB 29 CCAACGTGTGATCTGCGCCACCGCGGACCATCGCGGCGCCAGCGCGCA 88
QY 188 AGAG--TAGTACTCTGTGGAGCAGTCAACCGTTGATTAACCTTTGACGCGCTCCCTG 244
DB 89 ACAGCGCCACCTACAGAGGTTGGCTGCGAAGCTGATTTGCGCGCTGCGCG 148
QY 245 CCATCAAGACGCTGACCATCAAGGTTGACAGATCTCAAGCTTTGCTCCCAAGAGA 304
DB 149 AGCTGGCGACCTGGCCAAATGTGCGCGGAGCAAGTATGCAATGCTCCCAAGCA 208
QY 305 TGAAGGCTAGTGTGCTTAAGTACCAAGCTGTCAATGAGCTCTCTGCGCAAAAG 364
DB 209 TCAACCAAGCAGACCTCTCAAGCTGCAAGCAGCTGGCGGAGCTGGCCAGCAGCATG 268
QY 365 AGACCGAAGCGGTGATCATCAACCATGAACTGACACCATGGAAGAGACCGCTTTCTTC 424

Db 269 ACCTCATGCGATCGATCATACCATGGACGACGACCCCTGGAAGAACCGCTACTTTT 328
QY 425 TCACCTCAGCGTGAAGCAAAACCTGTCTGCTTGTAGCGCATGCTCAGGCT 484
DB 329 TGAACCTCGTGAAGAACGACGACCAAGCCGATCGCTGTGGTTCATATGCGCCGCGCA 388
QY 485 CTTCATGAGTGTGATGGCCCATGATCTCTATTAAGCGCTGAATGTAGCATCAAGA 544
DB 389 CCGCATGTCCGCGAGGCGATGCTCAACCTGTACACACCGCTGGCGCAGCAACA 448
QY 545 AAGCCTCTACTAACAAGAGAGTGTGATTTGTATGACGATGATTCACGCGCCAGAG 604
DB 449 AGGACTCGCGCGGCAAGGCGCTGTGTGATGACATGACAGAGATTCAGTCCGGCGTG 508
QY 605 AAGCAGCAAGCTCAACACCGCGAGTCAATGATTTGCTTCGCCCAACAGAGTAA 664
DB 509 ACGTGAGCAAGTCAATCAACATCAAGACCAAGCCTTC--AAGACGCGCTGGGCGCG 565
QY 665 TCGGACAGTCTATTATGCAAGTGTGATTTGATTCATGATTCGCTGACCTCACACC 724
DB 566 TGGGCTGTGTGTGAGAGCAAGTGTGATGCTGCTGCGCGGCAAGCGCCACACG 625
QY 725 TTGCAAGTGAAGTTGATTAATTAACAAATGCAAGAACTCCGAGAGTGTATTTCTTAC 784
DB 626 TCAACTCGAGTTGACATCAAGCAGATCAGCAGGCTGCCAGGTGACATGCGCTACA 685
QY 785 CTCACCCGATGATATGATGATTTAGTAAATGACACCTTACGACGAGGCAAGGAA 844
DB 686 GCTATGGAAGTCAACCGACGACGCTTCAAGGCGCTGCAAGAAAGCGCGCAAGCGCG 745
QY 845 TCATCCATGACGAGTGGCAATGGGAACCCCTTCCCTTGACTCAATAATGCTTGA 904
DB 746 TGATCCATGCGCGGACCGGCAATGGCTCGGTTCGCGGGTGTGCGACCTGCGACG 805
QY 905 AAGCAGCAATCAAGCGTGTGCTGCAAGCTTGAAGTGGCAGAGTGTTCACCA 964
DB 806 ACGTCGCAAGAACGCGTGCAGATCATTCGTCACGTCACAGGAGGCGTTTCGTGC 865
QY 965 CCCAAGAGCTGAATGATGATTAAGAACTGTTGTGTGCTACAGAGTCTCAAC 1024
DB 866 TCGCTAAGCGGAGAGCGCCGACGACAAAGACGATGGTGTGCGCCAGCAGCTCAAC 925
QY 1025 CTCAAAAGCAGAGTGTCTTATTTAGTGGCTTCACCAAACTGATGATAGAGGCGCA 1084
DB 926 CGCAGAAAGCGCGATCTGCGGATGTGGCATGACCAAGACCGAGCAGCAAGAGC 985
QY 1085 TCAGAAATATCTTCTCCACCTATTAT 1111
DB 986 TGCAGCGCATTTTCTGGGAATGATGAT 1012

RESULT 9
ABZ80801
ID ABZ80801 standard; DNA; 1174 BP.
XX
AC ABZ80801;
XX
DT 30-MAY-2003 (first entry)
DT
XX
DE Erwinia carotovora L-asparaginase gene.
KW gene; ds; asparaginase; cytostatic; blood; acute lymphoblastic leukemia.
XX
OS Erwinia carotovora.
XX
FH Key
FT CDS
FT Location/Qualifiers
FT 48..1094
FT /*tag= a
FT /product= "L-asparaginase"
FT /EC number= 3.5.1.1
FT primer_bind
FT 48..72
FT /*tag= b

```

FT primer_bind /bound_moiety= "primer N204"
FT primer_bind 1078..1100 /tag= C
FT primer_bind /bound_moiety= "primer C204"
XX MO2003018742-A2.
XX 06-MAR-2003.
XX 21-AUG-2002; 2002WO-RU00405.
XX 22-AUG-2001; 2001RU-0123442.
XX 04-APR-2002; 2002RU-0108505.
XX (UCHR-) GOS UCHREZHENTIE NI INST BIOMED KHIMI.
XX (UCHR-) GOS UCHREZHENTIE TSENTR BIOINZHENERIYA.
XX Eldarov MA, Zhigun AA, Gervaziev YV, Aleksandrova SS, Bogush VG;
XX Sidorchuk KV, Sveshnikova EV, Borisova AA, Omeinyuk NM, Archakov AI;
XX Skryabin KG, Sokolov NM;
XX WI: 2003-268418/26.
XX P-PSDB; ABP98615.
XX
XX New isolated Erwinia carotovora L-asparaginase gene useful for
XX producing a L-asparaginase protein useful for treating malignant and
XX nonmalignant diseases of the blood.
XX
XX Claim 2; Page 22-23; 29pp; Russian.
XX
XX This sequence represents a DNA fragment coding for an Erwinia carotovora
XX L-asparaginase. The invention also includes a homogeneous preparation of
XX a mature form of L-asparaginase produced by expression of the DNA in a
XX bacterial strain and has low affinity for L-glutamine. The L-asparaginase
XX protein encoded by the DNA is useful for treating malignant and
XX non-malignant diseases of the blood (e.g. acute lymphoblastic leukemia).
XX
XX Sequence 1174 BP; 290 A; 284 C; 313 G; 287 T; 0 other;
XX
XX Query Match 17.68; Score 199; DB 25; Length 1174;
XX Best Local Similarity 53.08; Pred. No. 7.2e-52;
XX Matches 483; Conservative 0; Mismatches 410; Indels 18; Gaps 2;
XX
XX QY 202 GCTGACGACGATCAAGCTTTGATTAAGCTTCTTGACAGCCCTCCCTGCATCAACGACCTAGCC 261
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX DB 204 GCGGGTGGCTGGCGGCTGAGAGAGCGCTGATTCAGACGGTGGCTGAGTGAAGAGCTTGCC 263
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX QY 262 ACCATCAAGGGTGAACAGATCTCAAGCATTTGGCTCCCAAGAGATGACGGGTAAAGTGTGG 321
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX DB 264 AATATCAAGGTGAGCAGGTGTCAGCATCGGCGAGTGAATAATAGACCAAGCATGTGTG 323
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX QY 322 CTTAAGCTAGCCAGCGTGTCAATGAGCTCTCGCCCAAAAAGAGACCGAAGCGGTGATC 381
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX DB 324 TTAAGCGTGAAGCAAGCGGCTGAGAGAGCTGCTGACAGCGCATGTGTGAGGTGTGTC 383
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX QY 382 ATCAACCCATGAGACTGACACCATGGAAGAGACGCGTTTCTTCTCAACCTCAGCGTAAA 441
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX DB 384 ATTACGACGAGTACGATACCTGACGAAATCTCTATTCTTGAACCTGACCGTGA 443
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX QY 442 AGCCAAAACCTGCTGCTGTGAGCGCCATGGGTGCACAGCTCTTCATGATGAGTGTGAT 501
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX DB 444 AGCGCAAGCCAGTAGTGTCTTTTGGCATGTCGCGCAAGGCAATCCAGTG---CCGAC 500
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX QY 502 GCGCCCATGAATCTCTATACGCGGTGAATGTAGCGATCAACAAAGCCTTACTAGACAA 561
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX DB 501 GCGCCGATGAACCTTACGCTGACATAAAGTGGAGGAGTAATAAAGTCTCCGCGGTGCGC 560
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX QY 562 GAGTGTGATGTGTATGAAGATAGATTAACGCGCCGACAGAGAGACCAAGTGTAC 621
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX DB 561 GGTGTACTGTGCTGCTTGAAGCAGCAGCATGTGTTGCTCCGTTTCATGAGAAACCAAC 620
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX QY 622 ACCACCGAGTCAATGCAATTTGCTTCCGCCAAGCAGGCTAAATGCGACAGTCTATAT 681
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11

```

```

DB 621 GCTTTACGTTGGATACCTTTAAAGCGCCAGAAAGGTTATCTGGGCGTGAATTACGT 680
QY 682 GCGAAAGTCAGATTTTCAATCGATCAATCGCTGACCTCACACCTTCGAAAGTATGAT 741
DB 681 GACAAATATATCTATCCAGACCGCTGCTGATTAAGTTACACACAGCGGTCCGTTGAT 740
QY 742 ATTAGCAAAATCGAAGACTCCCGAGAGTGAATTTTACGCTCACCCCGATGAT 801
DB 741 GTGACCAACGTTGATTAAGCTGCGCAAGTGCACATTTATGTTATGATATCCAGATCCA 800
QY 802 GATGTTTATGATATGACAGCCCTTACAGGACAGGACCCAAAGATCATTCATGAGCATG 861
DB 801 GAATATATATATGATGACCGCTTATCAAAACAGCGCTAAAGGACATGTTATGCGGGCATG 860
QY 862 GCGAATGGAGACCTTTTCCCTTGTGACTCAAAATGCTCTGAAAAAGCAGCCAAATGAGC 921
DB 861 GTGACAGCAGCAGTATTCACAGCGGGGAGCGCTGATTCGTAAGGAGGAAAGCAGAGC 920
QY 922 GTAGTGTGCTGCTGAGAGCTTAGAGTGGGAGCTGTTCCACACCCCAAGAGCTGAGATG 981
DB 921 ATTGTTGTGCTGATGCTTCCAGCGCTGACAGCGGATACGCTTCCACCGGATGAG----- 975
QY 982 GATGATTAAGAACTTGTGTTGTGCTGACAGAGATCTCAACCCCTCAAAAAGCAGAGTG 1041
DB 976 -----GCCAACCCGGTCTGTGTTCCGATTTCTGAGTCCAGGAAATACAGTAT 1025
QY 1042 CTTCCTTATGATAGCCCTCCACCAAACTAGATAGAGAGCGGATCAAAAGATCTTCC 1101
DB 1026 TTACTGATGCTGGCGCTTACGAAAGCAGCAACCGAGCTGTGATTCAGATTCCTCC 1085
QY 1102 ACCCTTATATC 1112
DB 1086 GCTTATATATC 1096
XX
XX RESULT 10
XX ID ABX66196 standard; DNA; 345 BP.
XX AC ABX66196;
XX AC 07-MAY-2003 (first entry)
XX DT Helicobacter pylori selected interacting domain (SID) DNA #795.
XX DE Protein-protein interaction; ulcer; selected interacting domain;
XX KW SID; gene; ds.
XX OS Helicobacter pylori.
XX PN WO200266501-A2.
XX PD 29-AUG-2002.
XX PE 28-DEC-2001; 2001WO-EP15428.
XX PF 02-JAN-2001; 2001US-259302P.
XX PK (HYBR-) HYBRIGENICS.
XX PA (INSP-) INST PASTEUR.
XX PI Legrain P, Rain J, Colland F, De Reuse H, Labigne A;
XX PI WI: 2002-674910/72.
XX DR P-PSDB; AB051452.
XX DR
XX PT New complexes of protein-protein interactions in Helicobacter pylori,
XX useful for identifying modulating compounds for treating or preventing
XX ulcers in mammals -
XX PS Claim 7; Page 279; 642pp; English.
XX PS
XX CC The invention describes a complex of protein-protein interactions in

```

CC Helicobacter pylori selected from 421 complexes given in the
 CC specification. The complex of protein-protein interactions are useful
 CC for screening for agents which modulate the interaction of proteins.
 CC Modulating compounds which binds to a targeted bacterial protein may be
 CC used for treating or preventing ulcers in a human or animal. This
 CC sequence encodes a selected interacting domain (SID), identified via
 CC protein-protein interactions.

XX Sequence 345 BP; 90 A; 69 C; 104 G; 82 T; 0 other;

Query Match 9.6%; Score 108.8; DB 24; Length 345;
 Best Local Similarity 57.7%; Pred. No. 1.1e-23;

Matches 194; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 142 CTAGCCACAGGAGGACCATCGCTGCTGGGGATCTAGCGTCAAGATGAGTACTCT 201
 Db 2 CTGGGACAGGGGGAGCATTCAGAGGAGTGTGGACGGGAGTTGGTATATAG 61
 QY 202 GCTGGACAGTCCACCGTTGATACCTTCTGACCGCTCCCTGCCATCAGACCTAGCC 261
 Db 62 AGTGGTAGTTGGGATCAAGACCTTTTGAAGGCTATCCCTAGCTTAACAGACTGCT 121
 QY 262 ACCATCAAGGTGAACAGATCTCAAGATTGCTCCCAAGAGATGAGCGGTAAGGTGG 321
 Db 122 CGCATTCAGGGGAGGAGATTTCTAACATGGCTCACAAGACATGAATGAAGAGGTATGG 181
 QY 322 CTTAATACGCAAGCGTGTCAATGAGCTCTCCGCCCAAAAGACCGGATGATC 381
 Db 182 TTCAAGCTCCCAAAACCTGCCCAAGAAATGCTAGATGATAGCCGTATTCAGAGGTGTC 241
 QY 382 ATCAACCATGAGACTGACACATGAGAGACCGCTTCTTCCCAACCTCAGCGTAA 441
 Db 242 ATCAAGCATGGACGACGACCTTTAGAGAGAGAGCGCGATTTTAACTAGTTTACGC 301
 QY 442 AGCAAAACCTGTCTGCTGAGGCGGCATGCGT 477
 Db 302 TCCACAAACCGGTGCTGTGGTGGAGCGATGCGT 337

RESULT 11

AA521124
 ID AA521124 standard; cDNA; 543 BP.

XX AA521124;

XX 21-MAY-2002 (first entry)

DE cDNA encoding Tm 3.9/His tag fusion protein minus signal peptide.

KW Anti-freeze peptide; Tm 3.9; yellow mealworm beetle; TRP; cryoinjury;
 KW thermal hysteresis protein; cryosurgery; cold climatization; ss;
 KW hypothermic cell preservation; de-icing formulation; transgenic plant;
 KW transgenic animal; His tag.

XX OS Tenebrio molitor.

XX Key Location/Qualifiers

XX CDS 64..513
 FT /*tag= a

FT misc-feature 64..165
 FT /*product= "Tm 3.9 clone protein"

FT mat_peptide 166..510
 FT /*note= "Nucleotides encoding His tag"

XX /*tag= c

XX WO200194378-A1.

XX 13-DEC-2001.

XX 07-JUN-2001; 2001WO-US18532.

XX 08-JUN-2000; 2000US-210446P.

XX (UNIV) UNIV NEW YORK STATE RES FOUND.
 PA (HORN/) HORNWATH K L.
 PA (MYER/) MYERS K L.
 PA (EAST/) EASTON C M.

XX Horwath KL, Myers KL, Easton CM;

XX WPI, 2002-090137/12.

XX P-PSDB; AAU10053.

PT New cDNA polynucleotide encoding a thermal hysteresis protein which is
 PT a type III anti-freeze protein derived from the Tenebrionidea
 PT Superfamily, useful for providing antifreeze protection to improve the
 PT quality of food -

PS Claim 4; Page 345; 364pp; English.

CC This invention relates to a cDNA polynucleotide comprising a nucleotide
 CC sequence encoding a thermal hysteresis protein (TRP) which is a type III
 CC anti-freeze protein derived from the Tenebrionidea Superfamily. Anti-
 CC freeze proteins lower the freezing point of a solution without affecting
 CC the melting point of the solution. An activated anti-freeze protein may
 CC be incorporated into plant, produce or fish in an amount sufficient to
 CC provide antifreeze protection or in a region of a target tissue to
 CC provide antifreeze protection to limit tumour cell or target tissue
 CC cryoinjury during cryosurgery. The proteins of the invention may also be
 CC used in hypothermic solutions or bathing media to reduce cold damage in
 CC order to provide cryogenic or hypothermic preservation of cells and
 CC tissues. The proteins may be used as de-icing formulations or used on
 CC surfaces to reduce existing ice buildup or abate the formation of ice
 CC buildup on surfaces such as a road, aircraft, household products,
 CC machinery and plant surfaces or as a food product to improve the quality
 CC of food by abating freezing of solutions, freezer burn, or degradation
 CC due to cold storage. The polynucleotides for the activated protein can
 CC be used to create transgenic or gene-modified plants, crops, fish, or
 CC animals having greater tolerance to cold climatization. The Tm 12.86
 CC antibody/antiserum which is also used as a screening device to screen
 CC cDNA libraries in an expression system, including cross-species cDNA
 CC libraries to identify homologous sequences in other species. The
 CC present sequence represents a cDNA encoding the Tm 3.9 clone/His
 CC tag fusion protein minus the signal peptide of the invention. This cDNA
 CC was created to facilitate purification and to try to enhance the anti-
 CC freeze activity of the recombinant protein.

XX Sequence 543 BP; 164 A; 133 C; 144 G; 102 T; 0 other;

Query Match 9.1%; Score 103.6; DB 24; Length 543;

Best Local Similarity 88.9%; Pred. No. 5.9e-22;

Matches 112; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 ATGGGACAGAGCCATCATCATCATCATATAGAGGGGCGCTGTGGCGGGGACGACCAT 60
 Db 64 ATGGGACAGAGCCATCATCATCATCATATAGAGGGGCGCTGTGGCGGGGACGACCAT 123
 QY 61 ATGGGACAGATGACTGTGGAGACAGCAAAATGGGTGGGATCCAGCCCTCTGTTTGATG 120
 Db 124 ATGGGACAGATGACTGTGGAGACAGCAAAATGGGTGGGATCCAGCCCTCTGTTTGATG 183
 QY 121 GCTAAA 126
 Db 184 CAGAAA 189

RESULT 12

AA521116
 ID AA521116 standard; cDNA; 543 BP.

XX AA521116;

XX 21-MAY-2002 (first entry)

DE cDNA encoding Tm 2.2 clone/His tag fusion protein minus signal sequence.

XX Anti-freeze peptide: Tm 2.2; yellow mealworm beetle; THP; cryoinjury;
 KW thermal hysteresis protein; cryosurgery; cold climatisation; ss;
 KW hypothermic cell preservation; de-icing formulation; transgenic plant;
 KW transgenic animal; His tag.
 OS Tenebrio molitor.
 XX
 FH Key Location/Qualifiers
 FT CDS 64..513
 FT /*tag= a
 FT /product= "Tm 2.2 clone protein"
 FT /tag= b
 FT /note= "Nucleotides encoding His tag"
 FT mat_peptide 166..510
 FT /*tag= c
 FT
 FT WO200194378-A1.
 XX
 PD 13-DEC-2001.
 XX
 PF 07-JUN-2001; 2001WO-US18532.
 XX
 PR 08-JUN-2000; 2000US-210446P.
 XX
 XX (UYNY) UNIV NEW YORK STATE RES FOUND.
 PA (HORN/) HORWATH K L.
 PA (MYER/) MYERS K L.
 PA (EAST/) EASTON C M.
 XX
 PI Horwath KL, Myers KL, Easton CM;
 DR MPI, 2002-090137/12.
 DR P-PSDB; AAU10045.
 XX
 PT New cDNA polynucleotide encoding a thermal hysteresis protein which is
 PT a Type III anti-freeze protein derived from the Tenebrionidea
 PT Superfamily, useful for providing antifreeze protection to improve the
 PT quality of food -
 PT
 XX
 PS Claim 4; Fig 5.8; 364pp; English.
 XX
 CC This invention relates to a cDNA polynucleotide comprising a nucleotide
 CC sequence encoding a thermal hysteresis protein (THP) which is a Type III
 CC anti-freeze protein derived from the Tenebrionidea Superfamily. Anti-
 CC freeze proteins lower the freezing point of a solution without affecting
 CC the melting point of the solution. An activated anti-freeze protein may
 CC be incorporated into plant, produce or fish in an amount sufficient to
 CC provide antifreeze protection or in a region of a target tissue to
 CC provide antifreeze protein to limit tumour cell or target tissue
 CC cryoinjury during cryosurgery. The proteins of the invention may also be
 CC used in hypothermic solutions or bathing media to reduce cold damage in
 CC order to provide cryogenic or hypothermic preservation of cells and
 CC tissues. The proteins may be used as de-icing formulations or used on
 CC surfaces to reduce existing ice buildup or abate the formation of ice
 CC buildup on surfaces such as a road, aircraft, household products,
 CC machinery and plant surfaces or as a food product to improve the quality
 CC of food by abating freezing of solutions, freezer burn, or degradation
 CC due to cold storage. The polynucleotides for the activated protein can
 CC be used to create transgenic or gene-modified plants, crops, fish, or
 CC animals having greater tolerance to cold climatisation. The Tm 12.86
 CC antibody/antiserum which is also used as a screening device to screen
 CC cDNA libraries in an expression system, including cross-species cDNA
 CC libraries to identify homologous sequences in other species. The
 CC present sequence represents a cDNA encoding the Tm 2.2 clone/his
 CC tag fusion protein minus the signal peptide of the invention. This
 CC cDNA was created to try to enhance the anti-freeze activity of the
 CC recombinant Tm 2.2 protein.
 XX
 SO Sequence 543 BP; 163 A; 133 C; 144 G; 103 T; 0 other;
 Query Match 9.0%; Score 102; DB 24; Length 543;

Best Local Similarity 88.1%; Pred. No. 1.9e-21;
 Matches 111; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 1 ATGGGAGACGACCATATCATCATCATATAGACGGGCGTGTGGCGGCGGACGACAT 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 64 ATGGGAGACGACCATATCATCATCATATAGACGGGCGTGTGGCGGCGGACGACAT 123
 QY 61 ATGGCTAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 124 ATGGCTAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 183
 QY 121 GCTAAA 126
 ||||
 Db 184 CAGAAA 189
 RESULT 13
 AAS21118
 ID AAS21118 standard; cDNA; 543 BP.
 XX
 AC AAS21118;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE cDNA encoding Tm 2.3 clone/his tag fusion protein minus signal sequence.
 XX
 KW Anti-freeze peptide: Tm 2.3; yellow mealworm beetle; THP; cryoinjury;
 KW thermal hysteresis protein; cryosurgery; cold climatisation; ss;
 KW hypothermic cell preservation; de-icing formulation; transgenic plant;
 KW transgenic animal; His tag.
 XX
 OS Tenebrio molitor.
 XX
 FH Key Location/Qualifiers
 FT CDS 64..513
 FT /*tag= a
 FT /product= "Tm 2.3 clone protein"
 FT /tag= b
 FT /note= "Nucleotides encoding His tag"
 FT mat_peptide 166..513
 FT /*tag= c
 FT
 FT WO200194378-A1.
 XX
 PD 13-DEC-2001.
 XX
 PF 07-JUN-2001; 2001WO-US18532.
 XX
 PR 08-JUN-2000; 2000US-210446P.
 XX
 XX (UYNY) UNIV NEW YORK STATE RES FOUND.
 PA (HORN/) HORWATH K L.
 PA (MYER/) MYERS K L.
 PA (EAST/) EASTON C M.
 XX
 PI Horwath KL, Myers KL, Easton CM;
 DR MPI, 2002-090137/12.
 DR P-PSDB; AAU10047.
 XX
 PT New cDNA polynucleotide encoding a thermal hysteresis protein which is
 PT a Type III anti-freeze protein derived from the Tenebrionidea
 PT Superfamily, useful for providing antifreeze protection to improve the
 PT quality of food -
 PT
 XX
 PS Claim 4; Fig 5.10; 364pp; English.
 XX
 CC This invention relates to a cDNA polynucleotide comprising a nucleotide
 CC sequence encoding a thermal hysteresis protein (THP) which is a Type III
 CC anti-freeze protein derived from the Tenebrionidea Superfamily. Anti-
 CC freeze proteins lower the freezing point of a solution. An activated anti-freeze protein may

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2003, 16:18:41; Search time 95 Seconds
(without alignments)
5264.075 Million cell updates/sec

Title: US-09-937-982-3

Perfect score: 1133
Sequence: 1 atgggcagcagccatcatca.....aagaagggaatctcttcac 1133

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued_Patents_NA:*

- 1: /cgn2_6/p/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/p/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/p/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/p/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/p/ptodata/1/ina/PTCUS.COMB.seq:*
- 6: /cgn2_6/p/ptodata/1/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1129.8	99.7	1133	3	US-09-094-435-3 Sequence 3, Appl
2	296.4	26.2	1830121	4	US-09-557-884-1 Sequence 1, Appl
3	296.4	26.2	1830121	4	US-09-643-990A-1 Sequence 1, Appl
4	294	25.9	1848	1	US-08-447-422-15 Sequence 15, Appl
5	215.8	19.0	1014	5	PCT-US92-10421-1 Sequence 1, Appl
6	210.8	18.6	1017	4	US-08-050-482A-1 Sequence 5793, Ap
7	200.4	17.7	1182	4	US-09-252-981A-5793 Sequence 5877, Ap
8	200.4	17.7	1182	4	US-09-252-981A-5877 Sequence 2336, Ap
9	159.2	14.1	1125	4	US-09-328-352-2336 Sequence 5773, Ap
10	147.4	13.0	2748	4	US-09-252-991A-5837 Sequence 5837, Ap
11	101.4	8.9	552	4	US-09-230-078A-3 Sequence 1293, Ap
12	80	7.1	270	4	US-09-107-532A-1293 Sequence 265, App
13	76.6	6.8	1005	4	US-08-961-527-265 Sequence 11, Appl
14	74	6.5	2231	4	US-09-613-303-11 Sequence 397, App
15	72.6	6.4	366	4	US-09-134-001C-397 Sequence 54, Appl
16	71.4	6.3	981	4	US-09-613-303-54 Sequence 20, Appl
17	70.4	6.2	1230	4	US-09-613-303-20 Sequence 4, Appl
18	66.2	5.8	1847	4	US-09-207-388-4 Sequence 110, App
19	65.2	5.7	1989	4	US-09-198-723A-110 Sequence 111, App
20	64.4	5.7	2016	3	US-09-198-723A-111 Sequence 13, Appl
21	64.4	5.7	2016	3	US-09-198-723A-111 Sequence 10, Appl
22	63.8	5.6	1624	4	US-09-647-224A-13 Sequence 112, App
23	63.8	5.6	1707	4	US-09-207-388-10 Sequence 113, App
24	62.8	5.5	648	3	US-09-198-723A-113 Sequence 114, App
25	62.8	5.5	648	3	US-09-198-723A-114 Sequence 115, App
26	62.8	5.5	648	3	US-09-198-723A-114 Sequence 115, App
27	62.8	5.5	648	3	US-09-198-723A-115 Sequence 115, App

28	62.8	5.5	651	3	US-09-198-723A-92 Sequence 92, Appl
29	62.8	5.5	651	3	US-09-198-723A-93 Sequence 93, Appl
30	62.8	5.5	651	3	US-09-198-723A-94 Sequence 94, Appl
31	62.8	5.5	651	3	US-09-198-723A-95 Sequence 95, Appl
32	62.8	5.5	651	3	US-09-198-723A-96 Sequence 96, Appl
33	62.8	5.5	651	3	US-09-198-723A-97 Sequence 97, Appl
34	62.8	5.5	651	3	US-09-198-723A-98 Sequence 98, Appl
35	62.8	5.5	651	3	US-09-198-723A-99 Sequence 99, Appl
36	62.8	5.5	651	3	US-09-198-723A-100 Sequence 100, App
37	62.8	5.5	651	3	US-09-198-723A-101 Sequence 101, App
38	62.8	5.5	1998	3	US-09-198-723A-102 Sequence 102, App
39	62.8	5.5	1998	3	US-09-198-723A-103 Sequence 103, App
40	62.8	5.5	1998	3	US-09-198-723A-104 Sequence 104, App
41	62.8	5.5	1998	3	US-09-198-723A-105 Sequence 105, App
42	62.8	5.5	1998	3	US-09-198-723A-106 Sequence 106, App
43	62.8	5.5	1998	3	US-09-198-723A-107 Sequence 107, App
44	62.8	5.5	1998	3	US-09-198-723A-108 Sequence 108, App
45	62.8	5.5	1998	3	US-09-198-723A-109 Sequence 109, App

ALIGNMENTS

RESULT 1
US-09-094-435-3
Sequence 3, Application US/09094435
Patent No. 6251388
GENERAL INFORMATION:
APPLICANT: Donald L Durden
TITLE OF INVENTION: UTILIZATION OF MOLINELLA SUCCINOGENES
TITLE OF INVENTION: ASPARAGINASE IN THE TREATMENT OF HUMAN
NUMBER OF INVENTION: HEMATOLOGIC AND AUTOIMMUNE DISEASE
CORRESPONDENCE ADDRESSES: 3
ADDRESS: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/094,435
CLASSIFICATION:
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/049,085
FILING DATE: June 9, 1997
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 234/274
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1133 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-094-435-3
Query Match 99.7%; Score 1129.8; DB 3; Length 1133;

Db	802956	TACTGATGCAATTTGCTATTACCCATGGTAGCAATACATGGAAGAGCGCTTAATTTCT	803015
QY	426	CAACCTCAGCGTGAAGAAAGCCAAAACCTGTCTGCTTGTAGCGGCATGCTCCAGGCT	485
Db	803016	AGATTTAACGGTAAATATGTGAAAAACCGGTTGTTCTCGTTGGGGCAATGCGCTCCGCAAC	803075
QY	486	TTCCATGAGTGCCTATGGCCCCATGAATCTGTATTAACCCCGGAAATGTAGGATCAACAA	545
Db	803076	AGAAAAAGTCTATGGCCCATTAATCTTTACATGCTGTGTTGTCGACAGCAGACAA	803135
QY	546	AGCCTCTACTAACAAAGAGTGTGATTTGATGAACAGATGATTCACGCCCGCAGAA	605
Db	803136	AAATTCAGTAGTGCTGTGGTGTGTTTAGTGCAGATGAATATGAAATACATAGGTGCTCGGA	803195
QY	606	AGCGACCAAGCTCAACACCACCGCAGTCAATGCAATTCGTTGCTGGCCCAACAGGTAAT	665
Db	803196	TGTAAACAAAAACAGTACAGCCGACGTCAAAAGTTCCATTGCACCAATTAATGTTCT	803255
QY	666	CGGACACATCTATTATGCGAAAGTCAGATATTTCACTCAATCGTTGCACTCACACCT	725
Db	803256	AGGCTATTTTCATATACACAAAGTGAGCATATGACGTTCCCGAAGAACATATACAT	803315
QY	726	TGCAAGTAGTGTGATTAATAGCAAAATGGAAGAACTCCCGCAAGTCGATATCTTTAGC	785
Db	803316	CAACACTCGGTTTAACTGTAAGAAAAATTAATGATAGCCTACCCAAAGTGGGATTATTTATG	803375
QY	786	TCACCCCGATGATCTGATGTTTGTGTCATGACACCCCTTCAGCGAGCAAGCAAGGAAT	845
Db	803376	TTATTCAATGCACTGTGTCGAACCTTAAGCAATTAACGATTAATCAATGCGTGTATCAAGAGAT	803435
QY	846	CATTCATGCAAGGATGGGCATGCGAACCCCTTCCCTTGACTCAAAATGCTCTTAAAA	905
Db	803436	TGTATCTGCGAGAGTGTGGCAATGGAATGTTATGCTGCACACTTGATCGCTTAAGAAA	803495
QY	906	AGCACCCAATATAGCGGTGTGCTGCGTCGGAAGCTCTAGAGTGGGAGTGTCCACAC	965
Db	803496	ACCCCGAANAAGATGGGTGCTGTGTACTCCGTTCTTCTGTCGTACCAACGGGTTATACAC	803555
QY	966	CCAAAGAGCTGAAGTGGATGATTAAGAACTTGTTTGTGGCTACAGAGAGTCTCAACC	1025
Db	803556	TCGTGAGCGTGAAGTTGATGATTAATAATATGCTTTGAGCATACAGTACTTAATCC	803615
QY	1026	TCAAAAAGCCAGAGTCTCTTATGTTAGCCCTCACCAAACTAGATGATGAGAGCGAT	1085
Db	803616	ACAGAAAGCCCGGTGCTCTTGCAATTAAGTTTAACCTCAAACTAAGATCCMAAAGTAAT	803675
QY	1086	CCAAAAGATCTTCTCCACCTATTAA	1110
Db	803676	TCACAATATTTTCGAGAGACTTTCTAA	803700
RESULT 3			
US-09-643-990A-1			
Sequence 1, Application US/09643990A			
Patent No. 6528289			
GENERAL INFORMATION:			
APPLICANT: Robert D. Fleischmann			
Mark D. Adams			
Owen White			
Hamilton O. Smith			
J. Craig Venter			
TITLE OF INVENTION: The Nucleotide sequence of			
the Hemophilus influenzae Rd Genome, Fragments			
thereof, and Uses Thereof			
NUMBER OF SEQUENCES: 1			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Human Genome Sciences, Inc.			
STREET: 9410 Key West Avenue			
CITY: Rockville,			
STATE: MD			
COUNTRY: USA			
ZIP: 20850			

```

COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PBI86P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match      26.2%; Score 296.4; DB 4; Length 1830121;
Best Local Similarity 57.1%; Pred. No.2.9e-86;
Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

QY 126 ACCCAATGACTATCTAGCCACAGAGGAGCACCATCGCTGGGTTCGGGGGAATACGCT 185
DB 802722 ACCAAATTTCACAACTTGGGCAAGGGGTGTACCATTCGAGAAAGCGGCAAAAGTTCGCT 802781

QY 186 CAAGAGTAGTACTCTGCTGGAGCAGTACCGGTGTATAGCTCTCTTGAGCGCGTCCGTC 245
DB 802782 AAATTCCTCGTATTAAGGTGGCAATTTAGTATGTATCTTTAATTGAAGGCTGTACCGA 802841

QY 246 CATCAAGCACTAGCCACCATCAAGGATGAAAGATCTCAAGCATTTGGCTCCCAAGAGAT 305
DB 802842 AATGAAAAATATGGCCACATTTAAAGGTGAGCAAAATTTGAAAAATAGGTTCACAAGACAT 802901

QY 306 GACGGGTAGGTGTGGCTTAAACTGTACCAAGCGGTGTCAATGAGTCTCTCCGCCCAAAAAGA 365
DB 802902 GAATGACGAAGCTGTGGCTAAAACTGGCAAAAAGCCATCAATGCTC-----AATGTAAAG 802955

QY 366 GACCGAAGCGGTATCATCAACCATGAGAACTGACACATGAGAAAGACCGCTTCTCTCT 425
DB 802956 TACTGTATGATTTGTTCATTACCATGTATACATACATGAAAGAGACGGCTTATTTCTT 803015

QY 426 CAACCTCAGCGGTGAAAAAGCCAAAACCTGTCTGCTTTGTAGCGGCATGCGTCCAGGCTC 485
DB 803016 ACATTTAACCGTAAATGTGAAAAAACCGGTTGTCTCGTTGGGCAATGCGTCTGCAAC 803075

QY 486 TTTCATGATGCTGTGATGGCCCATGAAATGCTATATAGCGCGGAATGTAGCATCAACAA 545
DB 803076 AGAAAAATGTGATGGGCCCATTTAATATCTTTTACATGCTGTGCTGTGCGCAGACAGAA 803135

QY 546 AGCCTGTACTTAACAAAGAGTGGTGTATGTGATGAACGATGAGATTACGCGCGCCAGAGA 605
DB 803136 AAATCAAGTGGTGTGATGGTGTATTTAGTCGAATGAAATGAAAGTACTAGGTGCTCGCA 803195

QY 606 AGCGCAACAGCTCAACACCAACCGCATTCATGATTTGCTGGCCCAACACAGGTAAAT 665
DB 803196 TGTAAACAAAACCAAGTACGACCCAGTGGCAAAAGTTTCATTCACCAAAATTATGTTCTCT 803255

QY 666 CGGCACAGTCTATTATAGCAAAAGTCAAGTATTTCACTCAATCCGTTGCACTCAACCT 725
DB 803256 AGGCTATATTCATACAGCAAAATGGAGCTGTGAACGTTCCCGAGAAAGCAAAATCATACAT 803315

```

```

QY 726 TCAGAGTGGATTTGATATATAGCAAAATCGAAGAACTCCCAAGTGCATATCTTACGC 785
Db 803316 CAACACTCCGTTTAACGTAGAAAAATAGATAGCTACCAAAAGTGGGATTTATATGC 803375
QY 786 TCACCCCGATGATATCTGATGTTTATGTCATGACGCCCTTACAGCAGCAAGCAAGAT 845
Db 803376 TTATATCAATGACACCTGCTGCAACCATTAACGATTAATCTCAATGCTGCTATCAAGGAT 803435
QY 846 CACCATGACAGGATGAGCAATGGGAACCCCTTCCCTTGCTCAAAATGCTCTGAAAA 905
Db 803436 TGATATCGCAGGATTTGGCAATGAAAAATGTAATGCTGACACATTAAATGCTTGA AAA 803495
QY 906 AGACGCCAATTCAGCGCTAGTCTGCTCGAAGCTGAGATGGGAGTGGTTCACACAC 965
Db 803496 AGCGGCAAAAGATAGCTGCTGTTAGTCCGTTCTCTGCTGACCAAGGGTTTACAC 803555
QY 966 CCAAGAGGCTGAGTGGATGATTAAGAACTGGTTGGTGGCTACAGAGTCTCAACC 1025
Db 803556 TCCTGACGCTGAAAGTATGATATGTAATATGGCTTTGTAGCAACAGGTTTAAATCC 803615
QY 1026 TCAAAAAGCAGAGTCTCTTATGTTAGCCCTCACCAAACTAGTATGAGAGGCGAT 1085
Db 803616 ACAGAAAGCCCGGCTGCTGCTGCAATTAAGCTTAACCACTAAGATCCMAAAGTAAT 803675
QY 1086 CCAAGATCTCTCCACCTATTAA 1110
Db 803676 TCACAAATATTTGGAAGACTTCTAA 803700

RESULT 4
US-08-447-422-15
; Sequence 15, Application US/08447422
; Patent No. 5686579
; GENERAL INFORMATION:
; APPLICANT: SHAMI, Ezekiel Y.
; APPLICANT: ROTHSTEIN, Aser
; APPLICANT: RAMJESINGH, Mohanb
; TITLE OF INVENTION: Use of Antibody/Antigen Interactions To
; TITLE OF INVENTION: Protect or Modulate Biological Activity
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,422
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/081,410
; FILING DATE: 22-JUN-1993
; APPLICATION NUMBER: US 07/938,505
; FILING DATE: 31-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/205,748
; FILING DATE: 21-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17923/102 HYLI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 15:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1848 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-447-422-15

Query Match
Best Local Similarity 56.3%; Score 294; DB 1; Length 1848;
Matches 574; Conservative 0; Mismatches 440; Indels 6; Gaps 1;

QY 91 GGTGCGGATCCACGCGCTGTTTGAATGGCTTAACCCCAAGAGCATATCTTGCCACA 150
Db 835 GGTGCTGGTGGTGGTGGTGGGCGGATCTAGATCTTACCAATTAACCATTTTGAACAC 894
QY 151 GGAAGCAACATCGCTGTTGCGGGAAATAGCGTCAAGATAGTACTGCTGGAGCA 210
Db 895 GCGGAGACCAATGCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 954
QY 211 GTCACCGTGTATAGCTTCTTGGACGCGCTCCCTGCTCAACGACCTAGCCACCATG 270
Db 955 GTTGGCGTAAATAATCTGTTAATGCGTCCGCGCACTAAAGACATTCGGAACGTAAA 1014
QY 271 GGTGAACAGATCTCAACGATTCGCTCCCAAGATGACGGGTAAAGTGGCTTAACTA 330
Db 1015 GCGGAGCAGGTAGTGAATATCGCTCCAGACATGACATTAATGCTGCTGAGCTG 1074
QY 331 GCCAAGCGTGTCAATGATGCTTCGCGCAAAAGAGACGCAAGCGGTATCATCCCAT 390
Db 1075 GCGAAAAAATTA-----CACCGACTGCGATTAAGACGCGGCTTCGCTATTAACCA 1128
QY 391 GGAAGTACACCATGGAAGAGACCGCTTCTCTCACTCACTCACTCACTCACTCACTCA 450
Db 1129 GGTACGACACGATGGAAGAACTGCTTACTTCTGACCTGACGGAATGCAACAAA 1188
QY 451 CCGTGTGCTTGTAGGCGGCGATGCTGCAAGGCTCTTCAAGTGTGATGCGCCCATG 510
Db 1189 CCGGTGTGATGCTGCGCGCAATGCGTCCGTCACGCTGTGAGCGGAGCGGTCCATTC 1248
QY 511 AATCTGTATAGACCGGTGATGATGATGATCAAAAGCCCTCTCAAAAGAGTGGTG 570
Db 1249 AACCTGTATAGACCGGTGATGATGATGATGATGATGATGATGATGATGATGATG 1308
QY 571 ATTGTATAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630
Db 1309 GTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1368
QY 631 GTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 690
Db 1369 GTAGCGACCTTCAAGCTGTTAACTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1428
QY 691 GAGTATTTCACTCAATCCGCTTCACTCACTCACTCACTCACTCACTCACTCACTCACT 750
Db 1429 GACTACACGCTGATCCCGGCGACGTAAGCATACAGAGACAGCCATTCATGCTCTAAG 1488
QY 751 ATCGAAGATCCCGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 810
Db 1489 GTGAATGAATCCGAAAGTGGCAATGTTGTTAACTAGGATGATGATGATGATGATGATG 1548
QY 811 GTCAATGACGCGCTTCAAGGAGAGCAAGAAATCAATCAATCAATCAATCAATCAATCA 870
Db 1549 GCTAAGCAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1608
QY 871 AACCTTCCCTTGTACTCAAAATGCTCTTGAAGAAAGCAAGCAATCAAGGCTAGTCTG 930
Db 1609 AACCTGTATTAATCTGTGTGACACGCTGGGCGACCGCGGAAACCGGTACTGCAATC 1668
QY 931 GCTGAAGCTAGAGTGGGAGGAGTGTTCACACCAAGAGGCTGAAGTGAATGAAG 990
Db 1669 GTGCGTCTTCCCGGCTACGAGCGGCGCTACCACTGCGAGATGCGGATGATGATG 1728
QY 991 AAACCTGTTTGTGCTAGAGAGTGTCAACCTCAAAAGCAAGAGTGTCTTATG 1050

```



```

COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/050,482A
  FILING DATE: 25-Apr-1995
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: WO PCT/US92/10421
  FILING DATE: 04-DEC-1992
ATTORNEY/AGENT INFORMATION:
  NAME: Bent, Stephen A.
  REGISTRATION NUMBER: 29,768
  REFERENCE/DOCKET NUMBER: 023032/0106
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (202) 672-5300
  TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 1017 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
  NAME/KEY: CDS
  LOCATION: 1..1011
  SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-050-482A-1

Query Match      18.6%; Score 210.8; DB 4; Length 1017;
Best Local Similarity 52.9%; Pred. No. 8.8e-60;
Matches 524; Conservative 0; Mismatches 457; Indels 9; Gaps 3;

QY 128 CCCAAGGACTATTCCTAGCCACAGAGGACACATCGCTGTCGGGGGAATCTAGCGTCA 187
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 29 CCAAGGCTGTGATCTCTCCGACCGCGGACACATCGCGGGCGCGCCAGCGCGGCA 88
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 188 AGAG--TAGCTACTGCTGTGAGCAGTCAACCGTTGATGACTCTTGACGCGCTCCCTG 244
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 89 ACAGCGCGACCTACACAGCGTCCAGAGTGTGGCGTCGACAAAGCTGATGGCGGCTGCGG 148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 245 CCATCAACGACCTTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGTCCCAAGAGA 304
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 149 AGCTGGCGGACCTGGCCAAATGTGCGCGGAGCAGATGATGACAGATCGCTCCGAAAGCA 208
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 305 TGACGGGTAAGTGTGCTTAAGCTAAGCAAGGCTGTCANTGAGCTCTCGGCCCAAAAG 364
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 209 TCACCAAGACGACCTCTCTACAGCTGGGCAAGCGCGTGGCGGAGCTGGCCAGACGATG 268
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 365 AGACCGAAGCGGTGATCATCACCATGGAAGTGAACACATGGAAGAACCGCTTTCTTCC 424
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 269 ACCTGATGTCATCATCATCACCATGGAAGTGAAGAACCGCTTTCTTCTTCTTCTTCTT 328
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 425 TCACCTCAAGGCTGAAGCAAAACCTGTCTGCTTGTAGAGCGCATCGCTCCAGGCT 484
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 329 TGACCTCTGGGAAAAGACCGACAGCGGCTGCTGCTGCTGCTTCCATCGCGCGGCA 388
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 485 CTTCGATGATGCTGATGGCCCATGATCTCTAATAGCGCTGAATGTGAGCATCAACA 544
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 389 CCGCCATGTCTCGCGACGCGCATCTCACTCTGTACAAAGCGCGTGGCGGCGGCAACCA 448
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 545 AAGCCCTACTAACAAGAGAGTGTGATGTGTGAACGATGAGATTCACGCGCCAGAG 604
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 449 AGAGCTCGCGGCGGAGGCGTGTGTGATGACCATGACGAGATCCAGTCCGCGCGTG 508
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 605 AAGCGCAAGGCTCAACACCGGACGATCAATGATTTGGTGGCCACACAGGTAAGA 664
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 509 AGGTAGCAAGTGTGATCAACATCAAGACCGAAGCCTTC--AAGAGCGCTGGGCGCGCG 565

```

```

QY 665 TCGGCACTATCTATTATGSCAAAGTCAGTATTTCATCATCCGTTCCAGCTCACACC 724
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 566 TGGGATGTGTGTGAGAGGCAAGTCGTAAGTCTGCTGCGCTGCGGCAAGCGCCACACG 625
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 725 TTGCAAGTGAAGTTATATTACCAAAATGGAAGAACTCCCGAGTGCATTTATTCTTACG 784
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 626 TCAACTCCGAGTTGACATCAAGCAAGATGAGAGGCTGCCCAAGGTGACATCGCTTACA 685
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 785 CTCACCCCGATGATCTACTGTTTATGTCATGACAGCCCTTAGGCAAGGAGGCAAGAGA 844
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 686 GCTATGGAAGCTCAACCGACAGCGCTTACAGAGCCCTTGACACAGAGCGCGGATTCG 745
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 845 TCATCATGACGAGCTAGGCGCAATGGGAACCTTTCCTTTGACTCAAAATGCTTGA 904
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 746 TGATCATGTCGGGACCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 805
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 905 AAGCAGCAAAATCAGAGCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 806 AGCTGCGCAAGACGCGCTGCAATCATCTTCTGCTCCAGCTCAACAGAGGCGGATTCG 865
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 962 CCACCAAGAGGCTGATGATGATGAAGAACTTGTGTTGTGCTACAGAGACTCA 1021
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 866 TGCTGCTGTAAGCGCGGACGCGCGGACGACAAAGACGATGGGTGCTGCGCCACGACCTGA 925
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1022 ACCCTCAAAAGCCAGAGTGTCTTATGTAAGCCCTTACCAAAACTGATAGAGAG 1081
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 926 ACCGCGAAGAGGCGCGGCTCTGCGGATGTGGCAATGACCAAGACCGACGACGAC 985
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1082 CGATCAAAAGATCTTCTCCACCTATTAAT 1111
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 986 AGCTGACGCGCATTTTCTGGGAATACTGAT 1015
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-09-252-991A-5793
Sequence 5793, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
  APPLICANT: Marc J. Rubenfield et al.
  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
  FILE REFERENCE: 107196.136
  CURRENT APPLICATION NUMBER: US/09/252,991A
  PRIOR FILING DATE: 1999-02-18
  PRIOR APPLICATION NUMBER: US 60/074,788
  PRIOR FILING DATE: 1998-07-27
  PRIOR APPLICATION NUMBER: US 60/094,190
  NUMBER OF SEQ ID NOS: 3142
  SEQ ID NO 5793
  LENGTH: 1182
  TYPE: DNA
  ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5793

Query Match      17.7%; Score 200.4; DB 4; Length 1182;
Best Local Similarity 52.5%; Pred. No. 2.8e-56;
Matches 513; Conservative 0; Mismatches 456; Indels 9; Gaps 3;

QY 128 CCCAAGGACTATTCCTAGCCACAGAGGACACATCGCTGTCGGGGGAATCTAGCGTCA 187
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 197 CCAAGGCTGTGATCTCTCCGACCGCGGACACATCGCGGGCGCGCCAGCGCGGCA 256
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 188 AGAG--TAGCTACTGCTGTGAGCAGTCAACCGTTGATGACTCTTGACGCGCTCCCTG 244
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 257 ACAGCGCGACCTTACACGCGCGGCAAGGTGCGGTCGACAGTGTGCGGAGGCTGCGG 316
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 245 CCATCAACGACCTTAGCCACCATCAAGGGTGAACAGATCTCAAGCATTTGGTCCCAAGAGA 304
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 317 AGCTCAAGGACATCGGCAAGCTCGGCGGCAAGAGTGTTCAGATCGCTCGGAAGCT 376
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 305 TGACGGGTAAGTGTGCTTAAGCTAAGCAAGGCTGTCAATGATGATCTCTGCGCCAAAAG 364

```

```

Db      377 TCACCAAGAGAACTGCTGGAACTGGGCAAGACCGTCCCAAGCTGGCGCATAGACGACG 436
QY      365 AGACGGAAGCCGTGATCATCAACCCATGGAACATGACATGGAAGAGACCGCTTCTTCC 424
Db      437 AGCTGAGGAGCTGCTGATCACCACCGTACCGACCCCTGGAAAGAGACGCTTCTTCC 496
QY      425 TCACCTCAAGGTGAAAAAGCCAAAAACCTGTGCTGCTTGTAGCGCCATGCTCCAGGCT 484
Db      437 TGACCTTGCTGAGACACACCGAAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556
QY      485 CTTCACAGAGTCTGATGAGCCCGCCATGATCTCTATATGCGCGTGAATGATGAGATCA 544
Db      557 CGGATGTCCTCCACCGCATGCTCAACCTGTACAAACCGCTGGCGGTGGCGGCGGAC 616
QY      545 AAGCCTCACTAACAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 604
Db      617 AGTGGGACGCGGCGGAGGCGGTGCTGATCAGATGACGACGAGATCTCTCGGCGCGG 676
QY      605 AAGCGACCAAGCTCAACACACCGCATGATGATGATGATGATGATGATGATGATGATGATG 664
Db      677 ACGGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 733
QY      665 TCGGACAGATCTATATGAGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 724
Db      734 TGGGATGCTGCTGAGGCGAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 793
QY      725 TTGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 784
Db      794 TGAACCTCGAGTTCGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 853
QY      785 CTCACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 844
Db      854 GGTACGAGCAAGTCAACGACGACCGCTACAGGCGCTGGCGCGCGCGCGCGCGCGCGCGCG 913
QY      845 TCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 904
Db      914 TCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 973
QY      905 AAGCAGCAATATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 961
Db      974 AACTGGGCAAGGAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1033
QY      962 CCACCCAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1021
Db      1034 TCTGGGCAAGCGGAGGAGCGGACGACGACGACGACGACGACGACGACGACGACGACGACG 1093
QY      1022 ACCCTCAAAAAGGAGGCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1081
Db      1094 ACCCGGAGAGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1153
QY      1082 CGATCCAAAAAGATCTCT 1099
Db      1154 AGCTGCAAGGATCTCT 1171

```

```

; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5877
Query Match      17.7%; Score 200.4; DB 4; Length 1335;
Best Local Similarity 52.5%; Pred. No. 3e-56;
Matches 513; Conservative 0; Mismatches 456; Indels 9; Gaps 3;

QY      128 CCCAAGTGAATATCTAGGACACAGAGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 187
Db      1211 CCACATGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1152
QY      188 AGAG--TACCTACTGCTGAGACATCAGCTGATGATGATGATGATGATGATGATGATGATG 244
Db      1151 ACAGGCCACACATACCGCGCGCAAGGTGCGGCTGACACGATGCTGCGGACGCTGCGCG 1092
QY      245 CCATCAAGACCTAGCCACCATCAAGAGGTGAAGATCAACATGCTGCTGCTGCTGCTGCTGCTG 304
Db      1091 AGCTCAAGGACATGCGCAAGCTCGCGCGCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1032
QY      305 TGAAGGATGAGTGTGCTTAATAGCCAAAGCGTGTCAATGATGATGATGATGATGATGATGATGATG 364
Db      1031 TCACCAAGGAGAACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 972
QY      365 AGACGCAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 424
Db      971 AGCTGACGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 912
QY      425 TCACCTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 484
Db      911 TGACCTGCTGAGCAACCGAAGAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 852
QY      485 CTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 544
Db      851 CCGGATGATGCTGCGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 792
QY      545 AAGCCTCTACTAACAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 604
Db      791 AGTGGGACGCGGCAAGGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 722
QY      605 AAGCGACCAAGTCAACACCGGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 664
Db      731 ACGGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 675
QY      665 TCGGACAGATCTATGAGCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 724
Db      674 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 615
QY      725 TTGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 784
Db      614 TGAACCTCGAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 555
QY      785 CTCACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 844
Db      554 GCTAGGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 495
QY      845 TCATCATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 904
Db      494 TCATCATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 435
QY      905 AAGCAGCAATATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 961
Db      434 AACTGCCAAGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 375
QY      962 CCACCCAGAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1021
Db      374 TCTGCGGCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 315
QY      1022 ACCCTCAAAAAGGAGGCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1081
Db      314 ACCCGAGAGGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 255

```

```

RESULT 8
US-09-252-991A-5877/c
; Sequence 5877, Application us/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5877

```

QY 1082 CGATCCAAAGATCTTCT 1099
 Db 254 AGCTGCAAGGATCTTCT 237

RESULT 9

US-09-328-352-2336
 ; Sequence 2336, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; TITLE OF INVENTION: BARMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 2336
 ; LENGTH: 1125
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-2336

Query Match 14.1%; Score 159.2; DB 4; Length 1125;
 Best Local Similarity 50.3%; Pred. No. 14e-42;

Matches 504; Conservative 0; Mismatches 483; Indels 15; Gaps 4;

QY 121 GCTAAACCCCAAGTACTATCTTACGACAGAGGACCATCGTGTGCGGGGAATCT 180
 Db 127 GCAAAAACATGTTGTTGTGCTGCTAGCGGCTACCATTCGCGGTGGGCGCAAGC 186
 QY 181 AGCGTCAGAGT---AGCTACTCTGCTGAGACGATCACCCTTGATTAAGCTTCTGACGC 237
 Db 187 TCAGAAATATGCACTTATATACGGCGGCAAAAGTTCAGTTGATGCTTATCAATGCA 246
 QY 238 GTCCCTGCCATCAACGACCTACACCATCAAGGGTGAAGATTCAGATTCAGATTCGCTCC 297
 Db 247 GTTCTCTCAATTCAGATTTGGCGATATCTGATATCCAGCAATTCAGATTCGCTCT 306
 QY 298 CAGAGATGACGGGTGCTGCTTAACTAAGCCAGCTGATGATGAGCTCTGCGCC 357
 Db 307 GAAAGTATCTGACAGAGATTTATACAAATGCTGCTCAAGTAAATGAACTTGTAA 366
 QY 358 CAAAAGAGACGGAAGCGGTGATCATACCCATGAGACCATGACCATGAGAGAGCCCT 417
 Db 367 AAGCCACTGTAAATGGCGTTGTGATTTACACAGGTACAGATCTTGAAGAAACAGCA 426
 QY 418 TTCTTCTCAACCTCACGCTGAAAAAGCCAAACCTGCTGCTGCTGAGGCGCATGCGT 477
 Db 427 TTTTCTTAATCTGTTGTCATACGATTAACCAATCGTACTGTTGGCTCAATGGCG 486
 QY 478 CCAGGCTTTCAGAGTGTGATGAGCCCATGAACTCTTAAACGCGGTGAATGAGG 537
 Db 487 CCAATCAATGCTTTTACAGAGATGTCACCTTAACCTTATAGCTGTTGCAATGGCC 546
 QY 538 ATCAACAAGACCTTACTACTACAAAGAGTGGTGAATGATGAAGATGAGATGACAGC 597
 Db 547 GCTTCTGATGATGAAAAAATAAGGCGTTATGGTTCTCATGAAGCACTCAATTTTGGCT 606
 QY 598 GCCAGAGAGGACCAAGCTCAACACCAAGCAGTCAATGCTTGGTTCGCCCAACACA 657
 Db 607 GCTGATGATGATTAAGGCAATTAACATCAATGCTTTTGAAGCC---AATGG 663
 QY 658 GGTAAATCGGACAGTCTATATGCAAGTGAAGTTCATCTCAATCCGTTGAGACT 717
 Db 664 GGTGCTTTGGTACACTTGTGAAGCAAAACCATTTGGTTAGACATCTGTTAAAGCT 723
 QY 718 CACACCTTGCAGAGTGTGATAT-----TAGCAAAATCGAAGAACCTCCAGAGTC 771
 Db 724 CATACCAATGCTTCAATTAATTAATGAAGTGAAGTGTCTTCCACAGTA 783
 QY 772 GATATCTTACGCTCAACCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 831

Db 784 CAAATGCTTATGATGCTGATCTATGCTTCTGATGCTTATGAAGCATATGAAGAAAGCT 843
 QY 832 GGAGCCAAAGATATCATCATGAGCATGCGCAATGGAA---CCCTTCCCTTGACT 888
 Db 844 GGCGCTAAAGCATTAATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 903
 QY 889 CAAATGCTTGAAGAAAGCAGCAATCAAGGCTGATGCTGCGTGAAGCTCTGAGAG 948
 Db 904 CCAATTTCAAAACCTTCATGACAAAAAGGATTTCAATTAATCCGTCATCGCGGCT 963
 QY 949 GCGAGTGTTCACACCAACCAAGAGCTGAAGTATGATGAAGAACTTGTGCTGCT 1008
 Db 964 CCACAGGTTTGTATTTAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1023
 QY 1009 ACAGAGCTCAACCTCAACCAAGAGAGTGTCTTATGATGATGATGATGATGATGATGAT 1068
 Db 1024 GCTATGATTTAAATCTCAAAAGGCGCTCTTCTGCGGCTGCGGCTGATCAATCAAAAGC 1083
 QY 1069 AGTATGAGAGGCGATCCAAAGATCTTCTCCACTATTAA 1110
 Db 1084 AAGCATGCGAAGAAATTCACAGCATGTTCTGCGAGTACTAA 1125

RESULT 10

US-09-252-991A-5773
 ; Sequence 5773, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 5773
 ; LENGTH: 2748
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-5773

Query Match 13.0%; Score 147.4; DB 4; Length 2748;
 Best Local Similarity 51.8%; Pred. No. 2e-38;
 Matches 384; Conservative 0; Mismatches 351; Indels 6; Gaps 2;

QY 362 AAGAGACCGAAGCGGTGATCATACCCATGGAATGACACCATGGAAGAGACCGCTTCT 421
 Db 3 AAGAGACCGAAGCGGTGATCATACCCATGGAATGACACCATGGAAGAGACCGCTTCT 62
 QY 422 TCCCTCAACCTCAAGCTGAAAAAGCCAAACCTGCTGCTGATGAGGCGCATGCTGAG 481
 Db 63 TCTGACCTCTGCTGAGACACACCAAGAGCCATCTGCTGCTGAGTGTGAGTGTGAGTGTGAG 122
 QY 482 GCTCTTCATGAGTGTGATGAGCCCATGAAATCTTATACGCGGTGATGATGATGATGATGAT 541
 Db 123 GCACCGCATGTGCGCGCAGCGAGATGCTCAACCTGATACAAAGCGGTGCGGTGCGCGG 182
 QY 542 ACAAGCTCTACTACAAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 601
 Db 183 ACAAGTGTGCGACGCGGCAAGGCGGTGCTGATCACCATGAAAGAGAGATCTCTCCGCC 242
 QY 602 GAGAGAGGACCAAGCTCAACACCAAGCGAGTCAATGATGATGATGATGATGATGATGATGAT 661
 Db 243 GCGAGCGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 299
 QY 662 AATGCGACAGCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 721
 Db 300 CGCTGGGATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 359

; LOCATION: (1)...(363)
US-09-613-303-11

Query Match 6.4%; Score 72.6; DB 4; Length 366;
Best Local Similarity 94.9%; Pred. No. 4.5e-14;
Matches 75; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ATGGGACAGCCATCATCATCATATGACAGCGGCTGTGCGCGCGGACCCAT 60
|||
Db 1 ATGGGACAGCCATCATCATCATATGACAGCGGCTGTGCGCGCGGACCCAT 60
OY 61 ATGGCTAGCATGACTGTG 79
|||
Db 61 ATGGCTAGCATGACTGTGAG 79

Search completed: September 4, 2003, 18:15:29
Job time : 102 secs

OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

NAME/KEY: misc_feature
LOCATION: (152530)...

Query Match 26.2% Score 296.4; DB 14; Length 1830121;
Best Local Similarity 57.1%; Pred. No. 5.3e-89;
Matches 562; Conservative 0; Mismatches 417; Indels 6; Gaps 1;

126 ACCCAAGTACTACTGACGACAGGAGGACGACATCGCTGGTTCGGGGAATCTAGCGT 185
1802722 ACCCAATATATACAACTTGGCAACGGGTGTGATACCTGACGAGGCGGCAAACTTGGT 802781
186 CAAGAGTACTACTGCTGGAGGAGGACGCTGTGATAGCTTCTTGGAGCGCTTCTGC 245
802782 AAATTCGCGTATTAAGCTGGACATTAATTAATTAATTAATTAATTAATTAATTAATTC 802841
246 CATCAGACCTAGCCACCATCAAGGTGAGACATCTCAAGTATGAGTCTCCCAAGAT 305
802842 AATGAAATATTTCCCAATTAAGGTGAGCAATATGTAATATAGTTCACAGACAT 802901
306 GACGGGTAAAGTGTGGCTTAACATGACCAAGCGTGTCAATGAGCTCTCGCCCAAAAGA 365
802902 GAATGAGAGTCTGGCTTAATGAGCAAAAGCCATCAATGCTC-----AATGTAAG 802955
366 GACCGAAGCCGTGATCATCAACCATGAGACATGACCAAGAGAGCGCTTCTTCT 425
802956 TACTGATGATTTGCTCATTAACCATGATGATGATGATGATGATGATGATGATGAT 803015
426 CAACCTAGGTTGAAAGCCCAAAACCTGTCTGCTTGTAGGCGCATGCTCAGGCTC 485
803016 AGATTTACCGTAAATGTAAGAAACCGGTTGCTGCTGGGCAATGCTCCTGCAAC 803075
486 TTCCATGATGCTGATGAGGCGCCATGATCTCTAAGCGCGTAATGAGCATCAACA 545
803076 AGAAAAAGTGTGATGAGGCGCCATTAATCTTTCATATGCTGTGCTGACAGCAG 803135
546 AGCTTACTAACAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 605
803136 AAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 803195
606 AGCGACCAAGCTCAACACGACGACATGATGATGATGATGATGATGATGATGATGAT 665
803196 TGTAAACCAAAACAGTACGACGACGATGATGATGATGATGATGATGATGATGAT 803255
666 CGGACAGTCTATTAATGCAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 725
803256 AGGCTATATTTCAACAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 803315
726 TGCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 785
803316 CAACACTCGTTTAAAGTAAATTAATGATGATGATGATGATGATGATGATGATGAT 803375
786 TCACCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 845
803376 TTATTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803435
846 CATTCATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 905
803436 TGTATTCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 905
906 AGCAGCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803495
803496 AGCCGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 965
966 CCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1025
803556 TCGTACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 803615
1026 TCAAAAAGCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1085
803616 ACAGAAAGCCGCTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1085
1086 CCAAAAGTCTTCTCCACCTATTA 1110

DB 803676 TCAACATATTTGCAAGACTTCTTA 803700

RESULT 3
US-09-882-227-83

Sequence 83, Application US/09882227
Publication No. US20030158396A1
GENERAL INFORMATION:
APPLICANT: Kleantous, Harold
APPLICANT: Al-Garawi, Amal
APPLICANT: Miller, Charles
APPLICANT: Tomb, Jean-Francois
TITLE OF INVENTION: Identification of Polynucleotides
TITLE OF INVENTION: Encoding No. US20030158396A1el Helicobacter Polyptides in t
FILE REFERENCE: 06132/047002
CURRENT APPLICATION NUMBER: US/09/882,227
CURRENT FILING DATE: 2001-06-15
PRIOR FILING DATE: 1997-07-29
NUMBER OF SEQ ID NOS: 638
SOFTWARE: FastSeq for Windows Version 4.0
LENGTH: 1196
TYPE: DNA
ORGANISM: Helicobacter pylori
NAME/KEY: CDS
LOCATION: (41)...(1132)
US-09-882-227-83

Query Match 22.9% Score 259.2; DB 12; Length 1196;
Best Local Similarity 55.5%; Pred. No. 3.3e-78;
Matches 548; Conservative 0; Mismatches 428; Indels 12; Gaps 2;

126 ACCCAAGTACTACTGACGACAGGAGGACGACATCGCTGGTTCGGGGAATCTAGCGT 185
157 ACCCAAGTACTACTGACGACAGGAGGACGACATCGCTGGTTCGGGGAATCTAGCGT 216
186 CAAGAGTACTACTGCTGGAGGAGGACGCTGTGATAGCTTCTTGGACCGCTTCTGC 245
217 TTGGGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 276
246 CATCAAGACCTTACCAACATCAAGGTGATGATGATGATGATGATGATGATGATGATGAT 305
277 TCTTAAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 336
306 GACGGGTAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 365
337 GAATGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 396
366 GACCGAAGCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 425
397 TATTTCAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 456
426 CAACCTCAGGTTAAAGCCAAACCTGTCTGCTTGTAGGCGCATGCTCCAGGCTC 485
457 AAATCTAGTTTAAAGCTCAACAAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 516
486 TTCCATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 545
517 TTCTTTGACGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 576
546 AGCTTCTACTAAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 605
577 AAAAAGTGTGAATTAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 636
606 AGCGAACAGCTCAACACGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 665
637 AGTGAATTAAGAGCACACCCACACCTTCACACCTTTAAAGCTTAATATACGCGGAT 696
666 CGGACAGTCTATTTAGCAAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 725

```

Db      697 AGGAGCGCTGATATATGCAAAAGCCGCTATTACATGACGCTTGAGAAAACACACAC 756
      726 TGCAGTGTGTTTATTTAGCAAAATGAA---GACTCCCGAGAGTCGATATTTCTTTA 782
      757 AGAGAGCGAATTTTCCCTTTTACACACTAAAACCCCTGCTTAAGTGGATATTTATTTA 816
      783 CGCTACCCCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
      817 CACGATGCTGGCAGACCCCTGATTTATTCACAGCAGCCTAACTGATGCAAAAG 876
      843 AATCATTCATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 902
      877 CATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 936
      903 AAGAGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 962
      937 AGAAGCGAGCCAAATGGGGGTGTTATTTGTTCTGACAGGATTAATAGCGGTGAGAT 996
      963 CACCCAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1022
      997 TA-----CTTACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1047
      1023 CCTCAAAAGCGAGTCTCTTATGTTAGCCCTCACCAAGCTAGTATAGAGAGC 1082
      1048 CCCCCAAAGAGTAGGGGTCTTTTACACTCGCTTTAACTAAACAAATATTAAGAAAA 1107
      1083 GATCAAAAGATCTTCCACCTATTAA 1110
      1108 AATCAAGAAATGTTGAAGATATTGA 1135

```

```

RESULT 4
US-09-842-628-1
: Sequence 1, Application US/09842628
: Patent No. US20020064862A1
: GENERAL INFORMATION:
: APPLICANT: ROBERTS, JOSEPH
: APPLICANT: MACALISTER, THOMAS W.
: APPLICANT: SETHURAMAN, NATARAJAN
: APPLICANT: FREEMAN, ABIE G.
: TITLE OF INVENTION: GENETICALLY ENGINEERED GLUTAMINASE AND ITS USE IN
: TITLE OF INVENTION: ANTIVIRAL AND ANTICANCER THERAPY
: FILE REFERENCE: 023032/0108
: CURRENT APPLICATION NUMBER: US/09/842,628
: CURRENT FILING DATE: 2001-04-27
: PRIOR APPLICATION NUMBER: 08/050,482
: PRIOR FILING DATE: 1995-04-25
: PRIOR APPLICATION NUMBER: PCT/US92/10421
: PRIOR FILING DATE: 1992-12-04
: PRIOR APPLICATION NUMBER: DE P 4140003.8
: PRIOR FILING DATE: 1991-12-04
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 1014
: TYPE: DNA
: ORGANISM: Pseudomonas sp.
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(1008)
US-09-842-628-1

```

```

Query Match      19.0%; Score 215.8; DB 9; Length 1014;
Best Local Similarity 52.6%; Pred. No. 2.9e-63;
Matches 519; Conservative 0; Mismatches 467; Indels 6; Gaps 2;
      128 CCCAAGTACTATCTTACACAGAGGACCATGCTGTGCGGGGAATCTACGCTCA 187
      29 CCAACGTGATCTGCGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 88
      188 AGAG---TAGCTACTCTGCTGAGAGAGTACCGGTGATAGAGTCTGTCAGCGGCTCTG 244

```

```

      89 ACAAGCGACCTACAGAGGCTGCCAAGGTGGCGTCCAGCAAGCTGATTTGCCGGCTCCG 148
      245 CCATCAACGACCTTAGCCACCATCATAGGTTAAACAGATTCACCATTTGGCTCCAGAGA 304
      149 AGCTGGCGGACCTGGCCAAATGTGCGGCGGAGGATGATGAGATTCGCTCCGAAAGA 208
      305 TGAAGGTAAGTGTGCTTAACTAGCCCAAGCGTCTCAATGATGATGATGATGATGATGATGAT 364
      209 TCACCAAGCAGACCTCTCAAGCTGCGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 268
      365 AGACCGAAGCGGTGATCATCATCCCATGAACTGACACCTGAGAGAGACCGCTTTCTCC 424
      269 ACGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 328
      425 TCACCTCAGCTGTAAGAAAGCAAGCAAGCGGATGCTGCTGAGGCGCATGCTGCAAGCT 484
      329 TGAACCTGCTGTAAGAAAGCAAGCAAGCGGATGCTGCTGAGGCGCATGCTGCAAGCT 388
      485 CTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 544
      389 CCGCATGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 448
      545 AAGCTCTACTAACAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 604
      449 AGGACTCGCGGCGGCAAGGCGGCTGCTGAGACCATGACAGAGATCCAGTCCGCGGCTG 508
      605 AAGCGACCAAGCTCAACACACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 664
      509 ACGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 565
      665 TCGGCAAGCTCTATTATGCAAAAGTGCATTTTCAATTCCTGCTGACCTTACACACC 724
      566 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 625
      725 TTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 784
      626 TCACTCGAGTTCGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 685
      785 CTCACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 844
      686 GGTATGCGAAGCTCAACCGCGGCTCAACAGGCGCTTGGCAAGAGAGCGGCGGCGGCGGCGGCGG 745
      845 TCATCATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 904
      746 TATTCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 805
      905 AAGCAGCCAAATCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 964
      806 AGCTGGCGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 865
      965 CCCAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1024
      866 TGGGTAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 925
      1025 CTCAAAAGCGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1084
      926 CGCAGAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 985
      1085 TCCAAAAGATCTTCTCCACCTATTAA 1111
      986 TGCAGCGCATTTCTGGAATACTGAT 1012

```

```

RESULT 5
US-09-876-348A-34
: Sequence 34, Application US/09876348A
: Patent No. US20020172951A1
: GENERAL INFORMATION:
: APPLICANT: Horvath, K. L. and Myers, K. L.
: TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
: TITLE OF INVENTION: Antifreeze Proteins and Method for Assaying Activity.
: FILE REFERENCE: RB-125-R1
: CURRENT APPLICATION NUMBER: US/09/876,348A

```

```

?
? CURRENT FILING DATE: 2001-08-09
? PRIOR APPLICATION NUMBER: 60/210,446
? PRIOR FILING DATE: 2000-06-08
? NUMBER OF SEQ ID NOS: 48
? SOFTWARE: Microsoft Word
? SEQ ID NO 34
? LENGTH: 543
? TYPE: DNA
? ORGANISM: Tenebrio molitor
? OTHER INFORMATION: His-tagged, Signal minus, Clone 3.9.
US-09-876-348A-34

```

Query Match	9.1%	Score 103.6;	DB 10;	Length 543;
Best Local Similarity	88.9%	Pred. No. 1.1e-24;		
Matches 112; Conservative	0;	Mismatches 14;	Indels 0;	Gaps 0

[illegible]

```

RESULT 6
US-09-876-796A-34
: Sequence 34, Application US/09876796A
: Patent No. US20020173024A1
: GENERAL INFORMATION:
: APPLICANT: Horwath, K. I. and Easton, C. M.
: TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
: FILE OF INVENTION: Antifreeze Proteins and Method for Assaying Activity
: FILE REFERENCE: RI-125-SEQO
: CURRENT APPLICATION NUMBER: US/09/876,796A
: CURRENT FILING DATE: 2001-06-07
: PRIOR APPLICATION NUMBER: 60/210,446
: PRIOR FILING DATE: 2000-06-08
: NUMBER OF SEQ ID NOS: 48
: SOFTWARE: Microsoft Word
: SEQ ID NO 34
: LENGTH: 543
: TYPE: DNA
: ORGANISM: Tenebrio molitor
: OTHER INFORMATION: His-tagged, Signal minus, Clone 3.9
US-09-876-796A-34

```

Query Match	9.1%	Score 103.6;	DB 10;	Length 543;
Best Local Similarity	88.9%;	Pred. No. 1.1e-24;		
Matches 112; Conservative	0;	Mismatches 14;	Indels 0;	Gaps 0

QY	1	ATGGGCGAGCGCATCATCATCATCATATACACAGCGCGCTGTCGCCCGCGGCGGACGCAT	60
Db	64	ATGGGCGAGCGCATCATCATCATCATCATCATCACACAGCGCGGCTGTCGCCCGCGGCGGACGCAT	120
QY	61	ATGGCGAGCATCATCTGCTGTCGACACGCAAAATGGTCCGCGATCCAGCGCCTCTGTTTGTATG	120
Db	124	ATGGCGTAGTACTGT	180
QY	121	GCTAAA	126
Db	184	CAGAAA	189

RESULT 7
US-09-876-348A-18
; Sequence 18, Application US/09876348A
; Patent No. US20020172951A1
; GENERAL INFORMATION:

```

; APPLICANT: Horwath, K. L. and Myers, K. L.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; TITLE OF INVENTION: Antifreeze Proteins and Method for Assaying Activity.
; FILE REFERENCE: RB-125-RI
; CURRENT APPLICATION NUMBER: US/09/87/6,348A
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/210,446
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 18
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Tenebrio molitor
; OTHER INFORMATION: His-tagged, signal minus, Clone 2.2
US-09-87/6-348A-18

```

Query Match	9.0%;	Score 102;	DB 10;	Length 543;
Best Local Similarity	88.1%;	Pred. No. 3.8e-24;		
Matches 11;	Conservative 0;	Mismatches 15;	Indels 0;	Gaps 0;

QY	ATGGGAGAGGCGCATATCATCATATATAGACAGCGGCTGTGGTCCGCGGGGAGGCAT	60
Db	64 ATGGGAGAGCGCATCATCATCATCAACAGACGGCGTGGTCCGCGGGGAGGCAT	123
QY	61 ATGGTAGCATCACTGTGGACAGCAATGGGTGGGATCCACAGGCTCTGTTTGTATG	120
Db	124 ATGGCTAGCATACATGGTGGACAGCAATGGGTCCGGATCCCTCACCGACAGACGATA	183
QY	121 GCTAAA	126
Db	184 CAGAAA	189

```

RESULT 8
US-09-876-348A-22
: Sequence 22: Application US/09876348A
: Patent No. US20020172951A1
: GENERAL INFORMATION:
: APPLICANT: Horwath, K. L. and Myers, K. L.
: TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
: TITLE OF INVENTION: Antifreeze Proteins and Method for Assaying Activity
: FILE REFERENCE: RB-125-RI
: CURRENT APPLICATION NUMBER: US/09/876,348A
: PRIORITY FILING DATE: 2001-08-09
: PRIOR APPLICATION NUMBER: 60/210,446
: PRIOR FILING DATE: 2000-06-08
: NUMBER OF SEQ ID NOS: 48
: SOFTWARE: Microsoft Word
: SEQ ID NO 22
: LENGTH: 543
: TYPE: DNA
: ORGANISM: Tenebrio molitor
: OTHER INFORMATION: His-tagged, Signal minus, Clone 2.3
US-09-876-348A-22

```

Query Match	9.0%;	Score 102;	DB 10;	Length 543;
Best Local Similarity	88.1%;	Pred. No. 3.8e-24;		
Matches 111;	Conservative	0;	Mismatches 15;	Indels 0;
				Gaps 0;

QY	1	ATGGGAGAGAGCCATATCATCATATCATATACACAGGCGCTGTCGCCCGCGGACGCCAT	60
Db	64	ATGGGAGAGAGCCATCATCATCATCATCATACACAGCAGGCGCTGTCGCCCGCGGACGCCAT	120
QY	61	ATGGGAGAGATACATCTGTCGACACCAATGGTCCGAGATCCAGGCGCTCTGTTTTGATG	120
Db	124	ATGGGAGAGATACATCTGTCGACACCAATGGTCCGAGATCCCTACCGAGACGACAGATA	180
QY	121	GCTAA	126
Db	184	CAGAA	189

QY 61 ATGGCTAGCATGCTGTGGACAGCAAAATGGTGGGATCCAGCCGCTCTGTTTGATG 120
|||||
Db 124 ATGGCTAGCATGCTGTGGACAGCAAAATGGTGGGATCCCTCACCAGCAAGATA 183
QY 121 GCTAAA 126
|||
Db 184 CAGAAA 189

RESULT 13

US-09-876-796A-30
; Sequence 30, Application US/09876796A
; Patent No. US20020173024A1
; GENERAL INFORMATION:
; APPLICANT: Horwath, K. L. and Easton, C. M.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; FILE REFERENCE: RB-125-SEQ
; CURRENT APPLICATION NUMBER: US/09/876,796A
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/210,446
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 30
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Tenebrio molitor
; OTHER INFORMATION: His-tagged, Signal minus, Clone 3.4
US-09-876-796A-30

Query Match
Best Local Similarity 9.0%; Score 102; DB 10; Length 543;
Matches 111; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 ATGGGCAGCAGCCATCATCATCATCATCATAGCAGCGGCTGTGGCGGGGAGCCAT 60
|||||
Db 64 ATGGGCAGCAGCCATCATCATCATCATCATAGCAGCGGCTGTGGCGGGGAGCCAT 123
QY 61 ATGGCTAGCATGCTGTGGACAGCAAAATGGTGGGATCCAGCCGCTCTGTTTGATG 120
|||||
Db 124 ATGGCTAGCATGCTGTGGACAGCAAAATGGTGGGATCCCTCACCAGCAAGATA 183
QY 121 GCTAAA 126
|||
Db 184 CAGAAA 189

RESULT 14

US-09-876-796A-38
; Sequence 38, Application US/09876796A
; Patent No. US20020173024A1
; GENERAL INFORMATION:
; APPLICANT: Horwath, K. L. and Easton, C. M.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Type III Tenebrio
; FILE REFERENCE: RB-125-SEQ
; CURRENT APPLICATION NUMBER: US/09/876,796A
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/210,446
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Microsoft Word
; SEQ ID NO 38
; LENGTH: 543
; TYPE: DNA
; ORGANISM: Tenebrio molitor
; OTHER INFORMATION: His-tagged, Signal minus, Clone 7.5
US-09-876-796A-38

Query Match
Best Local Similarity 9.0%; Score 102; DB 10; Length 543;
Matches 88.1%; Pred. No. 3.8e-24;

Matches 111; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 ATGGGCAGCAGCCATCATCATCATCATATGACAGCGGCTGTGGCGGGGAGCCAT 60
|||||
Db 64 ATGGGCAGCAGCCATCATCATCATCATATGACAGCGGCTGTGGCGGGGAGCCAT 123
QY 61 ATGGCTAGCATGCTGTGGACAGCAAAATGGTGGGATCCAGCCGCTCTGTTTGATG 120
|||||
Db 124 ATGGCTAGCATGCTGTGGACAGCAAAATGGTGGGATCCCTCACCAGCAAGATA 183
QY 121 GCTAAA 126
|||
Db 184 CAGAAA 189

RESULT 15

US-10-228-063-43
; Sequence 43, Application US/10228063
; Publication No. US20030135885A1
; GENERAL INFORMATION:
; APPLICANT: Lanahan, Mike
; TITLE OF INVENTION: Self-Processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: US/10/228,063
; CURRENT FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 1436
; TYPE: DNA
; ORGANISM: Thermotoga neopolitana
US-10-228-063-43

Query Match
Best Local Similarity 9.0%; Score 101.6; DB 12; Length 1436;
Matches 125; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 1 ATGGGCAGCAGCCATCATCATCATCATATGACAGCGGCTGTGGCGGGGAGCCAT 60
|||||
Db 1 ATGGGCAGCAGCCATCATCATCATCATATGACAGCGGCTGTGGCGGGGAGCCAT 60
QY 61 ATGGCTAGCATGCTGTGGACAGCAAAATGGTGGGATCCAGCCGCTCTGTTTGATG 120
|||||
Db 61 ATGGCTAGCATGCTGTGGACAGCAAAATGGTGGGATCCCTCACCAGCAAGATA 183
QY 121 GCTAAACCCAGTAGTACTTCTAGCCACAGAGCACCATGCC 164
|||
Db 121 GAGATCCGGAAGTGTGAGGCAAGAGAGTCCACCAACCC 164

Search completed: September 4, 2003, 19:20:51
Job time : 352 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2003, 16:12:46; Search time 2338 Seconds
(without alignments)
11778.003 Million cell updates/sec

Title: US-09-937-982-3

Perfect score: 1133
Sequence: 1 atggcagcagccatcatca.....aagaagggaatctcttcac 1133

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: EST:
1: em_estpa:*
2: em_estnum:*
3: em_estlin:*
4: em_estov:*
5: em_estov:*
6: em_estov:*
7: em_estov:*
8: em_estov:*
9: em_estov:*
10: em_estov:*
11: em_estov:*
12: em_estov:*
13: em_estov:*
14: em_estov:*
15: em_estov:*
16: em_estov:*
17: em_estov:*
18: em_estov:*
19: em_estov:*
20: em_estov:*
21: em_estov:*
22: em_estov:*
23: em_estov:*
24: em_estov:*
25: em_estov:*
26: em_estov:*
27: em_estov:*
28: em_estov:*
29: em_estov:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	155.2	13.7	983	BZ568327 pacs2-164
2	136	12.0	724	AC0990947 RfC01794
3	116.8	10.3	28	AC0500605 V36C6 mtn
4	110.8	9.8	757	BH376509 AG-ND-137

Result No.	Score	Query Length	DB ID	Description
5	108.6	9.6	843	AO500657
6	106.8	9.4	768	BZ570599
7	97.8	8.6	553	AO872693
8	97.8	8.6	553	AO872693
9	96.2	8.5	536	AO503225
10	96.2	8.5	536	AO503225
11	94.4	8.3	453	AO501722
12	94.4	8.3	453	AO501722
13	93	8.2	561	AO503362
14	89.6	7.9	663	BH371112
15	87	7.7	560	AO873892
16	79.2	7.0	840	BZ571432
17	75.6	6.7	644	AO157105
18	74	6.5	1378	BH770857
19	72.4	6.4	802	BZ571326
20	63.8	5.6	545	AO873272
21	60.6	5.3	781	CNS06FNL
22	60.6	5.3	871	BZ571362
23	55	4.9	427	AO991287
24	48	4.2	323	AO2049501
25	46.8	4.1	455	AO2049638
26	42.6	3.8	462	AO502215
27	42.6	3.8	559	BG672925
28	41.8	3.7	388	BZ570828
29	41.8	3.7	492	BZ570828
30	41.2	3.6	260	BZ570253
31	41.2	3.6	405	BZ570253
32	41.2	3.6	473	BZ570253
33	41.2	3.6	516	BZ570253
34	41.2	3.6	554	BZ570253
35	41.2	3.6	554	BZ570253
36	41.2	3.6	559	BZ570253
37	41.2	3.6	574	BZ570253
38	41.2	3.6	588	BZ570253
39	41.2	3.6	594	BZ570253
40	41.2	3.6	605	BZ570253
41	41.2	3.6	619	BZ570253
42	41.2	3.6	622	BZ570253
43	41.2	3.6	623	BZ570253
44	41.2	3.6	623	BZ570253
45	41.2	3.6	626	BZ570253

ALIGNMENTS

RESULT 1
BZ568327/c 983 bp DNA linear GSS 17-DEC-2002
pacs2-164_7476.y2 pacs2-164 Pseudomonas aeruginosa genomic clone
DEFINITION
pacs2-164_7476, genomic survey sequence.

ACCESSION
BZ568327
VERSION
BZ568327.1 GI:27201144

KEYWORDS
Pseudomonas aeruginosa
ORGANISM
Pseudomonas aeruginosa

REFERENCE
1 (bases 1 to 983)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence Variation among Multiple Isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press

JOURNAL
COMMENT
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: ckraymond@u.washington.edu
Class: Shotgun.
Location/Qualifiers

FEATURES

RESULT 2 LOCUS	DEFINITION	ACCESSION VERSION KEYWORDS	SOURCE ORGANISM	REFERENCE AUTHORS	TITLE	JOURNAL	MEDLINE
AO990947/c	AO990947 724 bp DNA linear GSS 14-AUG-2000	AO990947 R6C01794 Photorhabdus luminescens strain W14 M13 library sequence. AO990947 AO990947.1 GI:9649541 GSS.	Photorhabdus luminescens Photorhabdus luminescens Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Photorhabdus.	1 (bases 1 to 724) french-Constant, R.H., Waterfield, N., Burland, V., Perna, N.T., Daborn, P.J., Bowen, D. and Blatter, F.R.	A genomic sample sequence of the entomopathogenic bacterium Photorhabdus luminescens W14: potential implications for virulence	Appl. Environ. Microbiol. 66 (8), 3310-3329 (2000)	20378633

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycete

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 766)	Saccharomycetales: Saccharomycetaceae: Saccharomyces.			
Ros-Macdonald, P., Roemer, T., Coelho, P. S. R., Agarwal, S., Kumar, A., deesEtages, S. A., Cheung, K. -H., Sheehan, A., Symoniatelis, D., Jansen, R., Umansky, L., Heidtman, W., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roeder, G. S., and Snyder, M.		Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption	Unpublished	Contact: Kumar A Michael Snyder, Dept. of Mol. Cell. and Dev. Biology Yale University P.O. Box 208103, New Haven, CT 06520-8103, USA Tel: 203 432 9949 Fax: 203 432 6161 Email: anuj.kumar@yale.edu te of mtu-3xHA/lacZ insertion. Seq primer: GGCCCTCTTCTTGTGGAAGTAC Class: transposon-tagged.

```

FEATURES
    SOURCE
        Location/Qualifiers
            1..766
                /organism="Saccharomyces cerevisiae"
                /mol_type="genomic DNA"
                /db_xref="taxon:4932"
                /db_host="E. coli"
                /clone_lib="mtm-3kHA/lacZ Insertion Library"
                /note="vector: pHS56-Sal; A yeast genomic DNA library
                (lacking mitochondrial DNA) was prepared in pHS56-Sal;
                genomic DNA was size-fractionated (DNA of roughly 2-3 kb
                in length) prior to cloning. This library was
                subsequently mutagenized with a mtm-3kHA/lacZ
                multitransposon containing lacZ, URA3, and tet resistance.
    BASE COUNT
        223 a      175 c      168 g      198 t      2 others

```

	ORIGIN	
Query Match	Score 116.8; DB 28;	Length 766;
Best Local Similarity	Pred. No.2.9e-23;	
Matches 289; Conservative	Mismatches 287;	Indels 0; Gaps 0

[illegible]

Db 148 TTGGACACGAAATGCATGCCACCTTTAGTATCATTCAGAGGGGATGAACAGGGAT 83

Qy 662 AATCGGCACAGTCTATTATGCGAAAGTCAGATT 697

Db 88 ATTAGCTTACTTTCAAAATGATGACGTGAGATTTT 53

RESULT 4
 BH376509/c
 LOCUS
 DEFINITION BH376509 757 bp DNA linear GSS 10-DEC-26
 AG-ND-137P15, TF 1 ND-TAM Anopheles gambiae genomic clone
 AG-ND-137P15, genomic survey sequence.
 ACCESSION BH376509
 VERSION BH376509.1 GI:17322651
 KEYWORDS GSS.
 SOURCE Anopheles gambiae (African malaria mosquito)
 ORGANISM Anopheles gambiae

REFERENCE AUTHORS	TITLE	JOURNAL MEDLINE PUBMED COMMENT
1 (bases 1 to 757) Hong,S.S., Hogan,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J., F. 'C., Huff,E.R., Cantille,J.L., Black,K.K., Zhang,H.-B., Gardner,M.J. and Collins,F.H.	Construction of a BAC library and generation of BAC end sequence-tagged connectors for genome sequencing of the African malaria mosquito <i>Anopheles gambiae</i>	MOJ. genet. Genomics 268 (6), 720-728 (2003)
	22542063	
	12655398	
	Other_GSSs: AG-ND-137P15..TR.1	

Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel.: 301 838 0208
Fax: 301 838 3543
Email: bjoftus@tigr.org

This clone is from an *A. gambiae* BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from *A. gambiae* PST strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.
Seq primer: M13 For
Class: BAC ends.

FEATURES	source	Location/Qualifiers
		1..757
		/organism="Anopheles gambiae"
		/mol_type="genomic DNA"
		/strain="PES7"
		/db_xref="taxon:7165"
		/clone="AG-ND-13/P15"
		/clone_11b="ND-TAM"
		/note="Vector: pECBAC1; Site_1: HindIII"
BASE COUNT	182 a	196 c 192 g 187 t
ORIGIN		
Query Match	9.8%;	Score 110.8; DB 28; Length 757;
Best Local Similarity	50.4%;	Pred. No. 1.8e-21;
Matches 271; Conservative	0;	Mismatches 267; Indels 0; Gaps
OY	204	TGGACAGTCAACCGTTGATTAAGCTCTTCGACCGCTCCCTGCATCAACGACCTAGCCAC 2
DB	643	TTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1111
		TGCACCCATCTCTGCGCATGATCTCTTGCTACGATGGTGGCTTATGTAGGCCAAGTACGCCAA 5
OY	264	CATCAAGGATGAACGATTCATCAAGCATTTGGCTCCCAACAGATGACGGTAAAGTGTTGACT 3
DB	583	GATCCAGGTAAACAACCTTATTCACAGCTTCATCGGACTACATTCAGTGGAGGCTTGAGCCGTGGGT 5

OY 324 TAAACATGACCAAGCGTGTCAATGAGCTCTCCGCCAAAAAGACCGGATCAT 383
 Db 523 CCAGCTCAGTCTGCTGTGCAAGAGCGTATCAACAGTCACTGCTGGCGTATCGT 464
 OY 384 CACCCATGGAACATGACACCATGAGAGACCGCTTCTCTCCATCAACCTCAGCGTAAAG 443
 Db 463 TTCTCATGCGACAGATGCTGAGAGAAACAGCTTTTGGCTGTGATGACGTAAGTC 404
 OY 444 CCAAAAACCTGTCTGCTGTGAGCGCCGATGCTCCAGGCTTTCATGATGCTGATG 503
 Db 403 CAATTAACCAATGCTGATGAGCGGACGCAACCGCTCGGTTTCAAGATTGACGG 344
 OY 504 CCCCATGATCTCTATTAACCGCGGATGATGATGATCAACAAGCTTACTTAACAAG 563
 Db 343 TCCGCGCACTGTGCTTATGCTGCGCATGCTGTATCCATCCCAAGATCGTGG 284
 OY 564 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 623
 Db 283 CGCATGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 224
 OY 624 CACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 683
 Db 223 TCCCAACGTTGAACCTTTAAGTCTGATGATGATGATGATGATGATGATGATG 164
 OY 684 CAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 741
 Db 163 CAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 106

RESULT 5
 AOS00657/c 843 bp DNA linear GSS 29-APR-1999
 LOCUS v34f11 mtfn-3xHA/lacZ Insertion Library Saccharomyces cerevisiae
 DEFINITION genomic 5', genomic survey sequence.
 ACCESSION AOS00657
 VERSION AOS00657.1 GI:4706367
 KEYWORDS GSS.
 SOURCE Saccharomyces cerevisiae (baker's yeast)
 ORGANISM Saccharomyces cerevisiae

REFERENCE
 AUTHORS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 1 (bases 1 to 843)
 Ross-Macdonald, P., Roemer, T., Coelho, P. S. R., Agarwal, S., Kumar, A., deSeregas, S. A., Cheung, K. H., Sheehan, A., Symoniatidis, D., Jansen, R., Umansky, L., Heldman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Roeder, G. S., and Snyder, M.
 Large-Scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption

JOURNAL
 COMMENT Unpublished
 CONTACT: Kumar A
 MICHAEL SNYDER, Dept. of Mol. Cell. and Dev. Biology
 Yale University
 P.O. Box 208103, New Haven, CT 06520-8103, USA
 Tel: 203 432 9949
 Fax: 203 432 6161
 Email: anuj.kumar@yale.edu
 te of mtfn-3xHA/lacZ Insertion.
 Seq primer: GGCTTCTTCTTGTGGAATGAC
 Class: transposon-tagged.

FEATURES
 source
 1. 843
 /organism="Saccharomyces cerevisiae"
 /mol_type="genomic DNA"
 /db_xref="taxon:4932"
 /lab_host="E. coli"
 /clone_lib="mtfn-3xHA/lacZ Insertion Library"
 /note="Vector: PHSS6-Sal; A yeast genomic DNA library (lacking mitochondrial DNA) was prepared in PHSS6-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a mtfn-3xHA/lacZ multitransposon containing lacZ, URA3, and tet resistance."

BASE COUNT
 245 a 182 c 182 g 224 t 10 others

ORIGIN
 Query Match 9.6%; Score 108.6; DB 28; Length 843;
 Best Local Similarity 50.7%; Pred. No. 8.6e-21;
 Matches 255; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

OY 208 GCAGTACCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 267
 Db 679 GATTTAACCCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 620
 OY 268 AAGGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 327
 Db 619 GACTATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 560
 OY 328 CTAGCAAGGCTGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 387
 Db 559 TTGTATCAGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 500
 OY 388 CATGAACTGACACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 447
 Db 499 CATGAGACCGACACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 440
 OY 448 AAACCTGTGCTGCTTGTAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 507
 Db 439 AAGCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 380
 OY 508 ATGAATCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 567
 Db 379 ATGAATTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 320
 OY 568 GTGATG 627
 Db 319 ATGATCACTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 260
 OY 628 GCAGTACGATG 687
 Db 259 TCTTTGATACATG 200
 OY 688 GTGAGTATTTCACTCAATCCGT 710
 Db 199 GTGAGTATTTCACTCAATCCAGT 177

RESULT 6
 BZ570599/c 768 bp DNA linear GSS 17-DEC-2002
 LOCUS msh2_1446.y2 msh Pseudomonas aeruginosa genomic clone msh2_1446,
 DEFINITION genomic survey sequence.
 ACCESSION BZ570599
 VERSION BZ570599.1 GI:27205660
 KEYWORDS GSS.
 SOURCE Pseudomonas aeruginosa
 ORGANISM Pseudomonas aeruginosa
 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.

REFERENCE
 1 (bases 1 to 768)
 Spencer, D. H., Raymond, C. K., Smith, E. E., Sims, E. E., Hastings, M., Burns, J. L., Kaul, R., and Olsen, M. V.

JOURNAL
 COMMENT Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library J. Bacteriol., (2002) in press
 CONTACT: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 2062216954
 Fax: 2066857244
 Email: craymond@u.washington.edu
 Class: Shotgun.

FEATURES
 source
 1. 768
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"

```

/strain="MSH"
/db_xref="taxon:287"
/clone="msb2.1446"
/clone_lib="msb"
/Note="Environmental isolate. whole genomic shotgun
library."
BASE COUNT      116 a      246 c      239 g      165 t      2 others
ORIGIN

```

```

Query Match      9.48; Score 106.8; DB:29; Length 768;
Best Local Similarity 53.68; Pred. No. 2.8e-20;
Matches 287; Conservative 0; Mismatches 243; Indels 5; Gaps 3;

```

```

OY 331 GCCAAGCGTCAATGATGCTCTCCGCAAAAAGAGACGAAAGCCGTGATCATCACCAT 390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 610 GGCATACCGTGGCTGACGTGCGCCGATACGAAACAGTGGCGGATCGATGACACCCC 551
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 391 GGAATGACACCATGAGAGACCGCTTCTTCTTCACTACCTGAGGTGAAAGCCAAA 450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 550 GGTACGACACCTT-GAAGAGACCGCCCACTTCTGACCTGCTGAGCAGACCGGAGAG 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 451 CCTGCTGCTTGTAGCGCCATGCTCCAGCTCTTTCATGATGATGATGATGATGATG 510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 491 CCTATCGTGTGCTGCTGATGCGCCGCGGACCGCATGTCGCCGACGATGCTC 432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 511 AATCTTATAGCGCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 431 AACTCTACACGCGG-GGGCGTGGCGGGGACAACTGCGGACGCGGCAAGGCGGCTG 373
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 571 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 372 ATCACCATGACGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 313
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 631 GTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 690
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 312 ACCGAGGCGT---CAAGAGCCCGTGGGCGCGCTGGGATGCTGCTGCGGCGGACG 256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 691 GAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 750
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 255 TACTGTTCCGCGCGCGGCGGAGCGGACAGCGTGAATCGGATGCGATGCGATGCG 196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 751 ATGGAAGACTGCGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 810
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 195 ATCTCCGCGTGGCTGCGGATGATGATGATGATGATGATGATGATGATGATGATG 136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 811 GTCAATGACGCGCTTGGGCGGAGGCAAGGATGATGATGATGATGATGATGATGATG 865
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 135 TACAGGCGCTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 81

```

```

RESULT 7
AO872693/c      553 bp      DNA      linear      GSS 08-NOV-1999
LOCUS           V34E5 mtn-3xHA/lacZ Insertion Library, strain AB972 Saccharomyces
DEFINITION      cerevisiae genomic 5', genomic survey sequence.
ACCESSION       AO872693
VERSION         AO872693
KEYWORDS        GSS.
SOURCE          Saccharomyces cerevisiae (baker's yeast)
ORGANISM        Saccharomyces cerevisiae
REFERENCE       Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
                Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
AUTHORS        1 (bases 1 to 553)
                Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
                desRages, S.A., Cheung, K.-H., Sheehan, A., Symonistis, D., Jensen, R.,
                Umansky, L., Heidman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
                Lager, K., Miller, P., Roeder, G.S. and Snyder, M.
                Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
                Gene Disruption
                Unpublished
                Contact: Kumar A
                Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
                Yale University

```

```

P.O. Box 208103, New Haven, CT 06520-8103, USA
Tel: 203 432 9949
Fax: 203 432 6161
Email: anuj.kumar@yale.edu
te of mtn-3xHA/lacZ Insertion.
Seq primer: GGCCTTCTTCTTGGAGTAC
Class: transposon-tagged.
Location/Qualifiers
1..553
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="AB972 - tpi1 r(0) (S288c background)"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="mtn-3xHA/lacZ Insertion Library, strain AB972"
/Note="Vector: pHS56-Sal; A yeast genomic DNA library was
prepared in pHS56-Sal, genomic DNA was size-fractionated
(DNA of roughly 2-3 kb in length) prior to cloning. This
library was subsequently mutagenized with a mtn-3xHA/lacZ
multitransposon containing lacZ, URA3, and tet resistance."

```

FEATURES

```

source
1..553
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/strain="AB972 - tpi1 r(0) (S288c background)"
/db_xref="taxon:4932"
/lab_host="E. coli"
/clone_lib="mtn-3xHA/lacZ Insertion Library, strain AB972"
/Note="Vector: pHS56-Sal; A yeast genomic DNA library was
prepared in pHS56-Sal, genomic DNA was size-fractionated
(DNA of roughly 2-3 kb in length) prior to cloning. This
library was subsequently mutagenized with a mtn-3xHA/lacZ
multitransposon containing lacZ, URA3, and tet resistance."

```

```

BASE COUNT      152 a      127 c      124 g      150 t
ORIGIN

```

```

Query Match      8.68; Score 97.8; DB:28; Length 553;
Best Local Similarity 53.98; Pred. No. 1.1e-17;
Matches 201; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

```

```

OY 338 GTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 548 GTATCTCCGAGGACATGATGATGATGATGATGATGATGATGATGATGATGATG 489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 398 ACACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 488 ACACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 458 GCCTTGTAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 428 GTATCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 518 ATAAAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 368 ATCAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 309
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 578 TGAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 308 TAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 249
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 638 CATTGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 697
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 248 CATTCAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 698 TCACTCAATCGGT 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 188 ACTACCAACCACT 176

```

```

RESULT 8
AO503225/c      553 bp      DNA      linear      GSS 29-APR-1999
LOCUS           V34E05 mtn-3xHA/lacZ Insertion Library Saccharomyces cerevisiae
DEFINITION      genomic 5', genomic survey sequence.
ACCESSION       AO503225
VERSION         AO503225
KEYWORDS        GSS.
SOURCE          Saccharomyces cerevisiae (baker's yeast)
ORGANISM        Saccharomyces cerevisiae
REFERENCE       Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
                Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
AUTHORS        1 (bases 1 to 553)
                Ross-Macdonald, P., Roemer, T., Coelho, P.S.R., Agarwal, S., Kumar, A.,
                desRages, S.A., Cheung, K.-H., Sheehan, A., Symonistis, D., Jensen, R.,
                Umansky, L., Heidman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,
                Lager, K., Miller, P., Roeder, G.S. and Snyder, M.
                Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
                Gene Disruption
                Unpublished
                Contact: Kumar A
                Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
                Yale University

```


ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	COMMENT
AO873860								
AO873860.1	GI:6286104							
GSS.								
<i>Saccharomyces cerevisiae</i> (baker's yeast)								
<i>Saccharomyces cerevisiae</i>								
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;								
Saccharomycetales; Saccharomycetaceae; Saccharomyces.								
1 (bases 1 to 630)								
Ross-Macdonald, P., Roemer, T., Coelho, P. S. R., Aarval, S., Kumar, A.,								
desdanges, S. A., Cheung, K. H., Sheehan, A., Symoniatis, D., Hansen, R.,								
Umansky, L., Heidman, M., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R.,								
Hager, K., Miller, P., Roeder, G. S. and Snyder, M.								
Large-Scale Analysis of the Yeast Genome by Transposon Tagging and								
Gene Disruption								
Unpublished								
Contact: Kumar A								

QY	338	GTGTCAATGAGCTCTCTCGCCCAAAAAGAGACCGGAACCGGTATCATACACCATGGAAGT	397
DB	550	GTATCTCCGAGGACACTAGCTTCTGTGATGACTAGCGTGAGCCGTGTCACTCATCTGAGGACCG	491
QY	398	AACACATGGAAGAAGACCGCTTCTTCCTCCACCTCAGCTGAGGAAGGCCAAAACCTGTC	457
DB	490	ACACTATGGAGAGACGAGCTATCTCTTGTAGATTTCACCTAAATTCAGAGAACCAAGTAT	431
QY	458	GCCCTGTAGCGCCATGCGCTCCAGGCTCTTCCATGATGATGAGGCCCATGAATCTCT	517
DB	430	GTATCGCAGGCGGTATAGGCTCCAGGCACCTGACGCTCTGATGGCCCAATGAATTTAT	371
QY	518	ATPACGCCGTGAATGTAGCGATCAACAAACCTCTACTACAAAGGATGGTGATTTGTGA	577
DB	370	ATCAACCACTGTCTATATGCTGTCTTGAAATAATCACTGGGTGATGACGATGATCACTC	311
QY	578	TGAACGATGAGATTACGCGCGCCAGAGAAAGCGCAAGCTCAACACACCGGAGTCATG	637
DB	310	TAAACGATGTAATGGCTGTGGGTTTGGACAAGAAAATGAATGCAACTCTTTAGATA	251
QY	638	CATTTCCTTCGCCCAACAGCAAGTAATATGGCAGCTATATATGGCAAGTCAGATAT	697
DB	250	CATTCAAGACGAGATGAACAGGGAATTTATAGTTACTTTCAATGATATGAGGTGAGATTT	191
QY	698	TCAGTCAATCCGT	710

Db 190 ACTACCCACCAGT 178

RESULT	1.1
LOCUS	A0501722/c
DEFINITION	A0501722 453 bp DNA linear GSS 29-APR-1999
ACCESSION	VIGI min-3xHA/lacZ Insertion Library Saccharomyces cerevisiae genomic 5' , genomic survey sequence.
VERSION	A0501722
KEYWORDS	A0501722.1 GI:4707372
SOURCE	GSS.
ORGANISM	Saccharomyces cerevisiae (baker's yeast)
	Saccharomyces cerevisiae
	Saccharomycota; Ascomycota; Saccharomycotina; Saccharomycetes;

[illegible]

Qy	586	GAGATTACAGCGCCAGAGAACGACCGAACGTCACACACCGCAGCTCAATGCATTTGCT	645
Db	274	CGTATTGCTCTCGGGTTTTTGGACAACGAAAGAAATGCCACTCTTTAGATACATTCAGA	215
Qy	646	TCGCCAACACAGGTAAATCGCACAGCTATATTGCGAAGTCAGATTTCACTCAA	705
Db	214	GGCGATGAAACAGCGATATTAGTTACTTTCAATGATGACGTGGAGTTTACTACCCA	155
Qy	706	TCGCGT	710
Db	154	CCAGT	150
RESULT 14			
LOCUS	BH371112/c		
DEFINITION	AG-ND-137P15.TF ND-TAM Anopheles gambiae genomic clone AG-ND-137P15		
ACCESSION	BH371112	663 bp	linear
VERSION	BH371112		
KEYWORDS	BH371112.1 GI:17317237		
SOURCE	GSS.		
ORGANISM	Anopheles gambiae (African malaria mosquito)		
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anophelinae.		
AUTHORS	1 (bases 1 to 663) Hong,Y.S., Hogan,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J., Ren, C., Hiff,E.R., Carille,J.L., Black,K., Zhang,H.-B., Gardner,M.J. and Collins,F.H.		
TITLE	Construction of a BAC library and generation of BAC end sequence-tagged connectors for genome sequencing of the African malaria mosquito Anopheles gambiae		
JOURNAL	Mol. Genet. Genomics 268 (6), 720-728 (2003)		
MEDLINE	22542063		
PUBMED	12655398		
COMMENT	Other-GSSs: AG-ND-137P15.TR Contact: Brendan J Loftus Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0208 Fax: 301 838 3543 Email: b1o@fustel.igf.org This clone is from an A. gambiae BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from A. gambiae PEST strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest. Seq primer: M13 For Class: BAC ends.		
FEATURES			
SOURCE	Location/Qualifiers 1..663 /organism="Anopheles gambiae" /mol_type="genomic DNA" /strain="PEST" /db_xref="taxon:7165" /clone="AG-ND-137P15" /clone_lib="ND-TAM" /note="Vector: pCIBAC1; Site:1: HindIII"		
BASE COUNT	160 a 175 c 168 g 160 t		
ORIGIN			

Query Match	7.9%	Score 89.6;	DB 26;	Length 663;
Best Local Similarity	50.4%;	Pred. No. 3,4e-15;		
Matches 244;	Conservative	0;	Mismatches 239;	Indels 1;
			Gaps	1.

QY	259	GCCACATCAACGAGTGAACACATCCCAAGCATCGGGTAAAGGTG	318

QY	Db	RESULT 15	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
QY	589	GCCAAGATCCAGGTAACAACTTATCCACAGTTCATCGATGACATGAGGCGCTAGCCG	530											
QY	319	TGGCTTAAACTAGCCCAAGCGTGTCAATGATCTTCGCCCCAAAAGAGACGAAACCGTG	378											
Db	529	TGGGTCCACTACTACGTCTGTGTGCAAGCCGTATCAACAGTCAGAGTGTGCTGGCGTG	470											
QY	379	ATCATCACCATTGAACTGACACCATGGAAGACCGGCTTCTCTCTCAACCTCAACGGTG	438											
Db	469	ATCGTTTTCATAGGCAAGATACGCTGAGAGAAACAGCTTTTGGCTTGATTTGACAGTG	410											
QY	439	AAAAGCCAAAACCTGTCTGCCCTTGTAGCGCCATGCTCCAGGCTCTTTCATGATGCT	498											
Db	409	AAGTCCAAATTAACCAATGCTCTGTGATTTGGCGCAGACGCCAAGCCCTGGTTTCAACATTT	350											
QY	499	GATGGGCCCATCAATCTCATTAAGCCCGGATGTAGGATCAACAAAGCCCTCTATAC	558											
Db	349	GACGGTCCCGGCACTGTTTAAAGCTGTGCCATGCTGTTGATGCCCAATCCAGAGAT	290											
QY	559	AAAGAGATGATGATTTGTGTGAAGCATGATTCACGCCGCCAGAAAGCAGCAAGCTC	618											
Db	289	CGTGGCGCATCTGTGGCGATGAACACAGATCAACGGTACCAAGTATGTACCAAGAACA	230											
QY	619	AACACACCGGAGTAAATGATTTT-GCTTCGCCCAACACAGGTAAATCCGACAGTCTA	677											
Db	229	CACACTGCCAACGTTGAACTTTTAAAGTTGTGTGATGTTGGCTTTATTTGGAGGTCTA	170											
QY	678	TTATGGCAAAAGTCGAGTATTTCTCAATCCGTTGAGCTCAGACCTTGCAGAGTGT	737											
Db	169	CCGACAGCAAGGTCGTTTACCGCAGTAGGCGCATCCGCAATTTAACTTTCCATCAAGGC	110											
QY	738	TGAT 741												
Db	109	AGAT 106												
QY	AO873892	560 bp	DNA	linear	GSS 08-NOV-1999									
LOCUS	AO873892/c													
DEFINITION	VB335 mfn-3xHA/lacZ insertion library, strain AB972													
ACCESSION	AO873892													
VERSION	AO873892.1													
KEYWORDS	GI:6286136													
SOURCE	GSS.													
ORGANISM	Saccharomyces cerevisiae (baker's yeast)													
REFERENCE	Saccharomyces cerevisiae													
AUTHORS	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.													
	1 (bases 1 to 560)													
	Ross-Macdonald, P., Roemer, T., Coelho, P. S. R., Agarwal, S., Kumar, A., deEstiges, S. A., Cheung, K. H., Sheehan, A., Symoniatas, D., Jansen, R., Umanaky, L., Heidman, K., Nelson, K., Iwasaki, H., Kanada, D., Lugo, R., Hager, K., Miller, P., Koeder, G. S. and Snyder, M.													
	Large-Scale Analysis of the yeast Genome by Transposon Tagging and Gene Disruption													
	Unpublished													
	Contact: Kumar A													
	Michael Snyder, Dept. of Mol. Cell. and Dev. Biology													
	Yale University													
	P.O. Box 208103, New Haven, CT 06520-8103, USA													
	Tel: 203 432 9949				</									

FEATURES	SOURCE
Location/Qualifiers	1..560
	/organism="Saccharomyces cerevisiae"
	/mol_type="genomic DNA"
	/strain="AB972 - trp1 r(0) (S288C background)"
	/db_xref="taxon:4932"
	/lab_host="E. coli"
	/clone_lib="MTN-3kHA/lacZ Insertion Library, strain AB972"

/note="Vector: pHS6-Sal: A yeast genomic DNA library was prepared in pHS6-Sal; genomic DNA was size-fractionated (DNA of roughly 2-3 kb in length) prior to cloning. This library was subsequently mutagenized with a min-3xHA/lacZ minitransposon containing lacZ, URA3, and tet resistance.

BASE COUNT 155 a 127 c 119 g 158 t 1 others
ORIGIN

Query Match 7.7%; Score 87; DB 28; Length 560;

Best Local Similarity 54.0%; Pred. No. 1.8e-14;

Matches 177; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

```
QY 383 TCACCCATGGAACTGACACCATGAGAGACCGCTTCTCCTCAACCTCAGCGTGAAA 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 503 TCACTCATGGAGCCGACACTATGGAGAGACAGCTATTCTTAGATTGACCATANATT 444
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 443 GCCAAAAACCTGTCTGCCCTGTAGGCGCCATGCGTCAGAGCTCTTCATGAGTGTGATG 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 443 CAGAGAAGCCAGTATGTATCGACGCGATATGCTCCAGCCACTGCCAGCTCTGTGATG 384
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 503 GCCCATGATCTCTATATACGCCGTGATGTAGCGATCAACAAGCCTCTACTAACAAAG 562
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 383 GCCCATGATCTCTATATACGCCGTGATGTAGCGATCAACAAGCCTCTACTAACAAAG 562
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 563 GAGTGTGATGTGATGATGAGATTCACGCCGCCAGAGAGCGACCAAGCTCAACA 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 323 GCAGATGATCATCTTAACGATCGTATTCCTGCGGCTTGGACACGAAATGATG 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 623 CCACCGCAGTCATGCAATTTGCTTGGCCCAACACAGGTAAATCGGACAGTCTATTATG 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 CCAACTCTTTTATGATTCATTCAGAGCGGATGAAAGGATATTAGGTACTTTCAAAATG 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 683 GCAAGTCGAGTATTTCATCTCAATCCGT 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 203 ATGACGTGAGTTTACTACCACCACT 176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: September 4, 2003, 18:13:36
Job time : 2342 secs